

---

## Ⅲ章付録

# 指紋画像品質

Elham Tabassi  
Charles L. Wilson  
Craig I. Watson

**NISTIR 7151**  
**2004年8月**



---

要約.....	3
1 はじめに.....	3
1.1 用語.....	4
1.2 データセットと照合アルゴリズム.....	4
2 これまでの研究.....	6
3 指紋画像品質の定義.....	8
3.1 分類としての画像品質.....	12
4 指紋画像品質の測定.....	13
4.1 特徴抽出.....	13
4.1.1 MINDTCTへの入力.....	14
4.1.2 画像品質マップの生成.....	14
4.1.3 特徴点の品質の評価.....	15
5 指紋画像品質の評価.....	37
6 結論.....	68
参考文献.....	69
付録A.....	70
付録B.....	70

---

## 要約

このレポートでは、押捺指紋の品質の新しい定義を提案し、指紋の画像品質を測定するための詳細なアルゴリズムを示す。われわれは指紋画像品質を、照合アルゴリズムの適用前に照合性能を予測する目安であると定義する。つまり、品質のよい指紋画像を使用すると照合性能が向上し、反対に品質の悪い指紋を使用すると照合性能が低下することを意味している。また、指紋画像の品質評価に対する客観的な評価も実施した。この品質測定基準をCプログラミング言語で実装し、異なる操作設定で収集された20種類のライブスキャンおよび紙指紋のデータセットに対してテストした。この実装は、NISTの指紋解析ソフトウェアディストリビューションの一部として一般に公開されている（ただし輸出は規制されている）。

## 1 はじめに

NISTがPatriot Act(米国愛国者法)の一環として行った研究の結果、テストされた指紋システムの性能は、さまざまな要因のなかでも特に指紋画像品質によって大きく異なることがわかった[21、22、5、6]。この文書では、指紋照合性能の評価に使用できる、指紋画像品質の新しい測定基準を定義する。また、20種類のライブスキャンおよび紙指紋のデータセットに対するこの画像品質測定基準のテスト結果を示す。これらのデータセットについては、セクション1.2で説明する。さらに、商用指紋システムのベンダー8社が提供する14種類の指紋システムについて、照合性能の予測性能を評価した。NISTの指紋SDKテストの説明とデータセットの詳細な説明は、[6]で述べる。

理想的な状態で検出またはスキャンされた指紋画像には、はっきりと識別できる凹凸がある。自動指紋認識システムは、このような指紋画像については問題なく動作する。しかし、正確な指紋採取には独特の難しい側面がある[1]。皮膚の状態や不完全な採取のため、取り込まれた指紋画像は理想にほど遠い状態であることが多い。センサープレートが汚れていたり、接触の仕方にむらがあると[2]、サンプルの品質が低下し、画像処理中の特徴抽出誤差が生じるため、本人拒否率と他人受入率が増加する可能性がある。

自動指紋認識システムは、伝統的に指紋画像採取、特徴抽出、照合の3つのサブシステムで構成されてきた。画像採取では、人の指のライブスキャンか、紙(たとえば、指紋カード)の上に転写された人の指の押捺指紋のいずれかから、指紋のデジタル画像が取り込まれる。特徴抽出は、照合を容易にするために、取り込んだ画像を何らかの領域に描写するプロセスである。照合では、指紋がデータベース内の対象(人)から採取されたものである可能性が算出される。システム全体の性能は、個々のサブセットがどれだけうまく機能するかによって変わる。

照合プロセスの前に、指紋の「品質」を評価するのが望ましい。これにより、品質の低い採取画像がユーザのデータベースに入れられる前に修正できる。指紋の品質(この文書で後で定義する意味での品質)が十分によく、データベース全体の完全性が向上すれば、高い照合性能を達成できる。

以上の理由から、われわれは指紋画像品質を照合プログラムの性能の予測材料として定義する。これは、品質のよい指紋画像を使用すると照合性能が向上し、反対に品質の悪い指紋を使用すると照合性能が低下することを意味している。指紋画像品質については現在および過去の研究成果がいくつかあるが[2、11、12、13、14、16]、われわれの知る限り、指紋画像品質が照合性能を予測する尺度として公式に定義されたことはない。照合性能の予測は、複数の指紋を融合する生体認証でも重要である。これは、融合時に画像品質の最も高い指紋により大きい重みを割り当てることができるためである。

この文書の残りの部分は、次のように構成されている。まず、実験に使用する枠組みとデータセットについて簡単に説明する。セクション2では、指紋画像品質に関する現在および過去の研究成果を概説する。セクション3では、一般的な広い意味での指紋画像品質を定義する。ここでは実装の詳細については取り上げない。この定義は指紋を対象にした定義であり実装であるが、正しい特徴ベクトルが定義され、測定されれば、この品質定義をほかの種類（顔や虹彩など）にも適用することができ、また適用すべきであると、われわれは考えている。セクション4では、指紋品質を測定するための実装について詳しく説明する。セクション5では、品質システムのテスト方法について説明する。セクション6では、研究の成果を要約する。付録Aでは、20のデータセットすべてと14の指紋照合システムすべてを使ってこの方法のために行った画像品質予測の結果を示す。

## 1.1 用語

バイOMETリックサンプルとは、認識システムに入力されるデータを抽出したものである。例として、画像（指紋、顔、虹彩）、発語シーケンス、任意の合成データなどがある。ギャラリー  $\Gamma$  は、生体認証システムの登録者の集合を示す。このなかに、1つの対象につき必ず1つのバイOMETリックサンプルが含まれる。登録者の数を示すときは  $G = |\Gamma|$  を使用する。同様に、プローブセット  $\Pi$  はシステムの正規ユーザの集合である。一致は、同じ個人のプローブとギャラリーサンプルの比較結果を表す。不一致は、異なる人のサンプル間で発生する。

## 1.2 データセットと照合アルゴリズム

NISTは、異なる操作設定で採取されたライブスキャンおよびスキャンされた紙指紋の一連のデータセットを取得した。

米国国務省のデータセット(DOS-C)と米国国土安全保障省のデータセット(DHS2-C)には、左手と右手の人差し指のライブスキャン平面押捺指紋が含まれている。これらは、1つの運用環境で取り込まれたものであり、平面押捺指紋対平面押捺指紋(P2P)の照合のリアルタイムデータに期待される結果をもたらす必要がある。

米国国土安全保障省の十指指紋票(DHS10)とテキサス州公安局(TXDPS)のデータセットは、主に転写された紙のセグメント化された平面および回転押捺指紋であり、米国国土安全保障省給付金(BEN)はライブスキャンのセグメント化された平面および回転押捺指紋である。これら3つのデータセットは、平面押捺指紋対回転押捺指紋(P2R)の照合を評価するのに最も便利である。DHS10とTXDPSは、現在利用可能な指紋データベースの例であり、BENは品質をある程度制御できる新しいライブスキャン方法で取り込まれたデータの代表である。

US-VISITのPoint of Entry(VISIT\_POE)データとBio-Visa Application(VISIT\_POE\_BVA)データには、左手と右手の人差し指のライブスキャン平面押捺指紋が含まれている。これらは、1つの運用環境で取り込まれ、US-VISITプログラムに現在取り込まれている画像データを表している。BVAデータは国

---

務省の領事館がビザの申請者から取り込んだものであり、POEは国土安全保障省が米国への通関手続き場所で海外旅行者から取り込んだものである。VISIT\_POEはPOEで取り込まれた同一人物の2つのデータを比較し、VISIT\_POE\_BVAはPOE画像とBVA画像を比較する。

これらのデータセットのそれぞれから、6,000人のサンプルをランダムに選んだ。テストの開始後にDHS2-Cが整理統合されたため、統合後の最終的なデータセットのサイズは6,000ではなく5,888だった。統合のプロセスについては、[22]で詳しく説明されている。BEN、DHS10、およびTXDPSの4本指のストラップは、[10]で説明されているように、すべてNISTのセグメント化プログラムを使ってセグメント化されている。

NISTは、1対1のSDK(ソフトウェア開発キット)ベースのCOTS指紋照合システムのテストを実施し、1対1照合の精度を評価した[6]。8つの商用指紋ベンダーが各社のシステム(ベンダーによっては複数のシステム)を評価用に提出した。難易度の異なる20種類の1本指のデータセットについて、計14個のSDKベースの照合アプリケーションをテストした。画像品質評価システムを開発およびテストするため、SDKベンダーの類似度得点を使用した。このレポートでは、各SDKベンダーをアルファベット文字で表す。表1に、SDKを表す文字と対応するベンダー名を示す。

SDKを表す文字	ベンダー名
A	非公開
B	非公開
C	NEC
D	Cogent Systems, Inc.
E	非公開
F	Cogent Systems, Inc.
G	SAGEM Morpho, Inc.
H	NEC
I	Cogent Systems, Inc.
J	SAGEM Morpho, Inc.
K	Neurotechnologija Ltd.
L	非公開
VTB	NIST

表1. SDKを表す文字と対応するSDKベンダー名<sup>1</sup>

## 2 これまでの研究

生体認証に関する論文の多くは、指紋画像品質の評価に関する問題を扱っている。Nill & Bouzas [11]は、通常の場合で採取された場面のデジタル画像のパワースペクトルに基づいた客観的な画像品質を提案している。彼らのシステムは、デジタル画像の品質を評価するために設計されたものであり、指紋にも適用できる。Bolle, et al [12]は、品質測定基準として、無指向性領域に対する指向性領域の割合を使用した。Shen et al [13]は、凹凸パターンが明確に繰り返されるブロックを品質のよいブロックとして識別するため、画像のサブブロックにガボールフィルタを適用した。[12]と[13]は、いずれも局所的な方向性の情報のみを使用し、大域的な均一性や連続性に関する情報は無視している。局所的な隣接領域に含まれる凹凸の構成は、入力される指紋画像の品質によって異なるため、必ずしも明確な正弦波の凹凸が観察されるとは限らない。領域マスクをより正確に区別するには、大域的な特徴が必要である。どちらの論文も、提案する品質分析アルゴリズムの性能をテストするために主観的な方法を使用している。

Ratha and Bolle [14]は、指紋画像をWSQで圧縮するウェーブレット圧縮指紋画像による画像品質評価方法を提案した。しかし、彼らの品質測定基準の評価も主観的な内容である。

<sup>1</sup> これらのテストは、合衆国法典第8編第1732条で成文化されたBorder Security Act(国境保安法)第303条に従って、米国国土安全保障省のために実施された。このレポートで特定された具体的なハードウェアおよびソフトウェア製品は、この文書で説明する性能評価を実施するための技術の開発を適切に支援することを目的として使用された。したがって、NISTによる推薦または保証を意味するものではなく、特定された製品または装置が、その目的に関して得られる最善のものであ

ると意味しているわけでもない。

Hong et al[15]は、凹凸パターンをサイン波としてモデル化し、サイン波の振幅、周波数、および分散を計算することによって指紋の品質を決定した。彼らは、指紋の領域を回復可能な領域と回復不可能な領域に分類している。回復可能なブロックの比率が指定されたしきい値より小さい場合は、入力された指紋を品質が悪いものとして拒否する。

Lim and Yau[16]は、指紋画像の品質と妥当性を空間領域内で評価するためのアルゴリズムを提案している。彼らは、画像ブロックのグレーレベルの階調の共分散行列から得られた固有値の割合によって凹凸パターンの繰り返しを検証している。また、自動指紋識別システムによって指紋画像から抽出した特徴の品質を、提案する画像品質分析手順の性能をテストするためのベンチマークとして使用している。彼らのシステムは、品質の悪い指紋画像や無効な指紋画像を特定することを目的としている。

Yao et al[2]は、指紋画像に信頼できる方向性の情報がどれだけ含まれるかを測定する尺度として品質Qを算出している。彼らは、高い性能が得られたことを示すためにROC曲線を使用した。提案する品質分析アルゴリズムの性能のテストにはROC曲線を使用しなかった。

われわれの知る限り、バイOMETリックサンプルの品質を認識性能の予測材料として公式に定義した例はなく、順位付けされたROCを使って測定された品質を評価した例もない。

### 3 指紋画像品質の定義

指紋とは、指先の表面にある皮膚紋理のパターンのことである。品質のよい指紋には識別可能なパターンと特徴があり、それらを使い、後で指紋どうしの照合の際に有効な特徴を抽出することができる。特徴点に基づく自動指紋照合アルゴリズムでは、2つの指紋(バイOMETリックサンプル) $x_{g(i)}$ および $x_{p(j)}$ の局所的な隆線の特性(特徴点)を比較する関数を使って実数値の類似度得点が生成される。ここで、添字 $g(i)$ は*i*番目のギャラリーを示し、 $p(j)$ は*j*番目のプローブを示し、 $s_{ij}$ は*j*番目のプローブに対して照合された*i*番目のギャラリーの得点である。

$$s_{ij} = F(x_{g(i)}, x_{p(j)}) \quad (\text{eq. 1})$$

真の(同一人物の)比較の類似度得点 $s_{ii}$ を一致得点と呼び、偽の(異なる人物の)比較の類似度得点 $s_{ij}$  ( $i \neq j$ )を不一致得点と呼ぶ。したがって、規模が*G*のギャラリーで、システムに登録された対象ごとにバイOMETリックサンプルが必ず1つだけあると仮定すると、各プローブ画像 $x_{p(i)}$ に対して1個の一致得点 $s_{ii}$ と*G*-1個の不一致得点 $s_{ij}$ ,  $i \neq j$ が存在する。 $s_m(x_i)$ はサンプル $x_{p(i)}$ の一致得点を示し、 $s_n(x_{ji})$ は $x_{p(i)}$ と $x_{g(j)}$ (ただし、 $i \neq j$ )の不一致得点を示すとする。

高い類似度得点は、サンプルが同じ個人のものである可能性が高いことを示すと解釈される。 $M(s_m)$ は一致得点の累積分布関数(CDF)を示し、 $N(s_n)$ は不一致得点のCDFを示すとする。検出誤りトレードオフ特性(DET)は、誤った一致率

$$FMR=1-N(s_n) \quad (\text{eq. 3})$$

に対する誤った不一致率

$$FNMR=M(s_m) \quad (\text{eq. 2})$$

を、 $s_m$ および $s_n$ のすべての値についてグラフ化したものである。DET(および同等のROC)は、検証システムの性能の最も一般的な表現である。

ここでは、指紋画像品質を、照合プログラムの性能を予測する目安であると定義する。先に進む前に、照合性能を定量化する必要がある。類似度得点は、予測される性能の最終的な表現である。これにより、裏付けとなる一致と不一致の分布とともに、サンプルが同一人物または異なる人物から採取された可能性が示される。一致と不一致の分布は、複雑な非線形アルゴリズムの結果であり、通常はランダムではなく、内部的なアルゴリズムやパラメータの設定方法に強く依存している。

図1に、ベンダーVTB[10, 6]のデータセットNIST-SD29[3]に含まれる216人の押捺指紋に関する一致得点と不一致得点のヒストグラムを示す。216個の一致得点と46,440個の不一致得点がある。

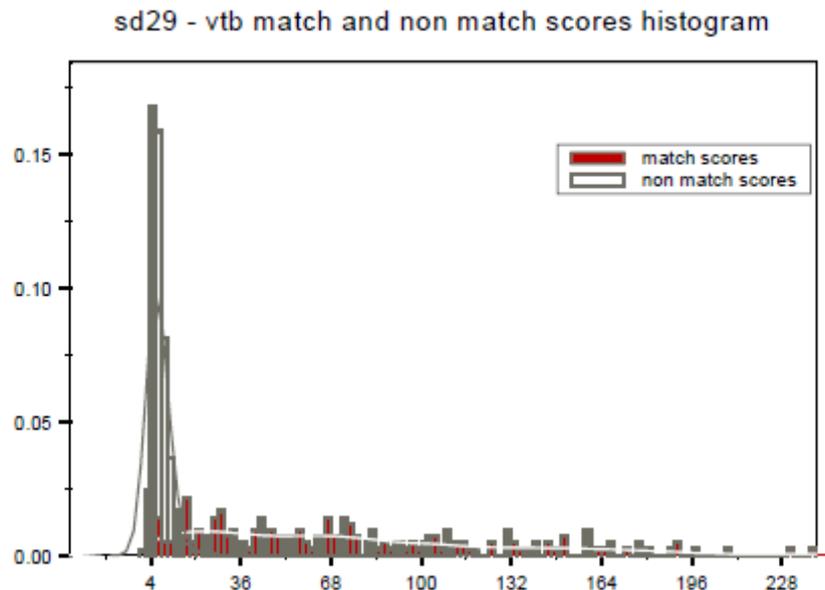


図1. SD29 VTBの一致得点と不一致得点のヒストグラム

一般に、不一致の分布より一致の分布の方が広がる。また、ほとんどの場合、2つの分布は重なり合う。一致と不一致の分布が重なり合うということは、あるサンプル $x_i$ の一致得点 $s_m(x_i)$ が不一致得点 $s_n(x_{ji})$ より小さい( $s_m(x_i) < s_n(x_{ji})$   $i \neq j$ )場合に、そのサンプルが誤って一致することを意味する。品質測定基準 $q$ によって照合性能を予測する場合、品質のよい指紋は一致得点の高い指紋であり、不一致の分布から十分に分離している必要がある。同様に、品質の悪い指紋は一致得点の低い指紋であり、特にその一致得点が不一致得点と重なり合う領域に含まれる指紋である。

したがって、品質測定基準 $q$ は一致の分布 $M(s_m)$ が不一致の分布 $N(s_n)$ からどの程度分離しているかを示す必要がある。具体的には、以下を予測するためにバイオメトリックサンプル $x_i$ の品質 $q_i$ を定義する。

$$o(x_i) = \frac{s_m(x_i) - E[s_n(x_{ji})]}{\sigma(s_n(x_{ji}))} \quad \forall x_i \in \Gamma \text{ or } \Pi \quad (\text{eq. 4})$$

ここで、 $E[\cdot]$ は数学的期待値であり、 $\sigma(\cdot)$ はサンプル $x_i$ の標準偏差、 $s_m(x_i)$ は一致得点、 $s_n(x_{ji})$ は不一致得点である( $\forall j, i \neq j$ )。プローブサンプル $x_i$ とサンプルが $G$ 個ある内部ギャラリー(同じ対象(人)のサンプルが必ず1つだけ含まれる)を比較すると、 $G$ 個の得点のベクトル $s$ が得られる。ベクトル $s$ の要素のうち1つだけが $x_i$ の一致得点であり、ほかの $G-1$ は $x_i$ の不一致得点である。 $E[\cdot]$ は、ギャラリー内の $G-1$ 個の不一致項目に対するプローブサンプルの不一致得点全体の平均を計算することによって求められる。同様に、 $\sigma(\cdot)$ は $s$ の不一致要素のみから求められる標準偏差である。 $o(x_i)$ をサンプル $x_i$ の正規化一致得点と呼ぶ。基本的には、対象のバイオメトリックサンプルを、一致したとされるサンプルと比較し、次に不一致サンプルと比較し、追加の得点に基づいて素点を調整する。

図2に、データセットDOS-Cの右手人差し指について方程式4を使って生成されたSDKベンダー照合プログラムFの正規化一致得点のボックスプロットを示す。この図は、正しく一致する指紋と誤って一致する指紋を別々に示している。 $o(\cdot)$ の値が小さい指紋は誤って一致するのに対し、正しく一致する指紋は $o(\cdot)$ の値が大きい。

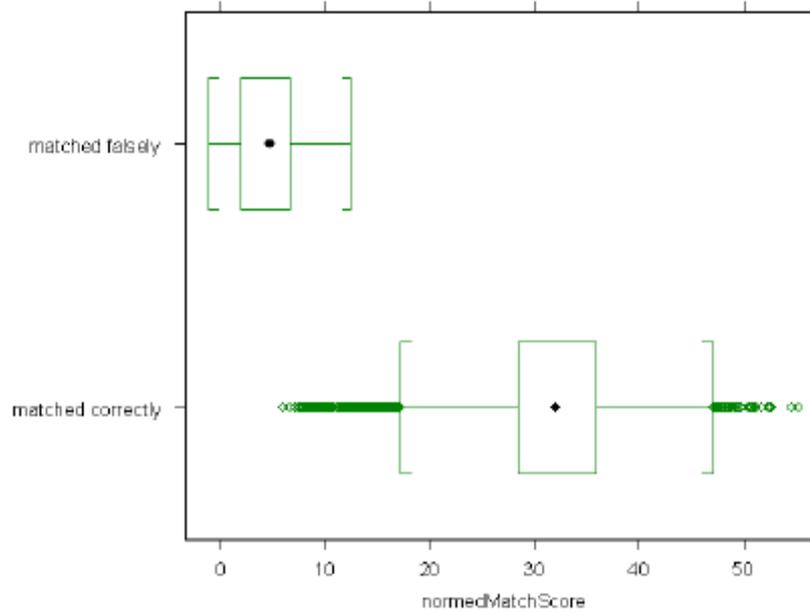


図2. 誤って一致する場合と正しく一致する場合を区別した $o(\cdot)$ のボックスプロット。ボックス内部の黒い点はデータセットDOS-Cの右手人差し指の指紋に関するベンダーFの正規化一致得点(方程式4)の中央値を表し、ボックスは得点の四分位範囲を示す。ひげは1.5(四分位範囲)に描画され、外れ値と見なされる範囲外の点は個別に描画されている。

各指紋について、その画像品質を正規化一致得点の予測材料として定義する。方程式1で定義した類似度得点と方程式4で定義した正規化一致得点は、プローブサンプルとギャラリーサンプルの両方の関数であるが、ここで定義する品質は、各サンプルについて別個に測定されたスカラー値である。したがって、次に定義されるペア品質 $q$ によってペア $(x_{gallery}, x_{probe})$ の認識性能を予測できるはずである。

$$q = H(q_{gallery}, q_{probe}) \quad (\text{eq. 5})$$

NIST[5, 6]の広範囲にわたるテストにより、品質の低いサンプルは、認識エラーを引き起こす要因となることがわかっている。つまり、 $H(\cdot)$ は個々の数値 $q_{probe}$ および $q_{gallery}$ の最小関数であるため、ペア品質は方程式6で定義される。この文書では、品質を方程式6で定義される $q$ とする。

$$q = \min(q_{gallery}, q_{probe}) \quad (\text{eq. 6})$$

ある操作設定で、登録したサンプルの品質が高いと仮定される場合、対象のバイOMETリックサンプル(プローブ)の品質を測定すれば、その正規化一致得点を十分に予測することができる。

ここでは、まず $x_i$ の適切な信号または画像の忠実度特性を含む特徴ベクトル $v_i$ を計算し、次に $v_i$ から $o(x_i)$ への(非線形の)写像を見つけることにより、バイオメトリックサンプル $x_i$ の(スカラー値の)品質 $q_i$ を測定する。数学的には次のように表される。

$$v_i = L(x_i) \quad (\text{eq. 7})$$

$$q_i = \tilde{o}(x_i) = I(v_i) \quad (\text{eq. 8})$$

関数 $L(\cdot)$ は、照合アルゴリズムの情報を伝達する $x_i$ の特性と特徴を計算することにより実現される。サンプル $x_i$ に $L(\cdot)$ を適用すると、 $n$ 次元の特徴ベクトル $v_i$ が得られる。指紋の場合は、このなかに凹凸の測定された明瞭度、画像のサイズ、および特徴点の数と品質の測定基準が含まれる。特徴ベクトルについては、セクション4で詳しく説明する。

関数 $I(\cdot)$ は、特徴ベクトルの空間 $v$ から正規化一致得点 $o(\cdot)$ への写像である。 $\tilde{o}(x_i)$ は、 $o(x_i)$ の予測値である。

### 3.1 分類としての画像品質

方程式4に示すように、バイオメトリックサンプルの品質を正規化一致得点の予測材料として定義した。方程式8は、変数のベクトル $v$ から応答関数 $o(\cdot)$ を求めるために、各種の回帰法を使用すべきであることを示している。しかし、セクション4で詳しく示すように、主に次の理由で各種の回帰法を試みても適切な方法は見つからなかった。その理由は、1つは、サンプル固有の測定基準(ベクトル $v$ の成分)は応答変数 $o(\cdot)$ に対して直線性を有していないため、非線形関数を見つける必要があることである。2つ目は、外れ値がデータに大きな影響を与えていることである。3つ目は、どの回帰法を試みても残差がガウス分布にならないことである。また、品質のレベル(最高、高、中、低など)がわかれば十分である。つまり、品質は $o(\cdot)$ (正規化一致得点)の予測材料として定義されるため、各サンプル $x_i$ の正確な値ではなく、 $o(x_i)$ の範囲(最高、高、中、低など)がわかれば十分である。これらの事実から、この問題を分類の観点から提示し直すことにする。つまり、サンプルの品質を $o(x_i)$ の瓶を予測する測定基準として定義する。方程式8の関数 $I(\cdot)$ は、基本的にサンプル $x_i$ の特徴ベクトル $v_i$ (方程式7)を品質値 $Q$ に写像する分類器である。これにより、 $o(x_i)$ は一般性を失わずに $k$ 番目の瓶に分類される( $Q = k$ )。つまり、サンプル $x_i$ の品質値 $q_i$ は、 $o(x_i)$ を分類するための瓶である。われわれの実装では、この分類器として人工ニューラルネットワーク[17]を採用している。

許容される瓶の数 $K$ は、粗さのパラメータであり、品質値とは相殺関係にある。たとえば、高/低レベル( $K = 2$ )は前述のような $K$ が無限に大きくなる連続的な場合より容易に実現できる。この文書では、 $K = 5$ とする。つまり、品質のレベルを不可(5)、可(4)、良(3)、優(2)、秀(1)の5種類とする。画像品質を5レベルに分けるのは、1つの妥協案である。NIST[6]の研究によれば、画像品質にひじょうに敏感な照合プログラムを完全に特徴付けるには8~10レベルが必要である。画像品質にあまり敏感でない照合プログラムであれば、画像品質を3レベルに分けても十分に特徴付けることができるだろう。ここでは、品質が中~高レベルの照合プログラムにとって適切な5レベルを選択した。 $o(\cdot)$ の瓶境界を定義するため、不規則に間隔をあけた $o(\cdot)$ の分位を使用した。たとえば、分位関数 $x$ は次に示すように経験的CDFの逆関数である。

$$\text{quantile}(p) = \min\{z \mid o(x) \geq \text{割合} p\}$$

$o(\cdot)$ の瓶境界を表2のように定義する。ここで、 $W(\cdot)$ は誤って一致するサンプルのCDFを示し、 $C$ は正しく一致するサンプルのCDFを示す。

Q	品質	範囲
5	不可	$[0, W^{-1}(0.75)]$
4	可	$[C^{-1}(0.75), C^{-1}(0.05)]$
3	良	$[C^{-1}(0.05), C^{-1}(0.2)]$
2	優	$[C^{-1}(0.2), C^{-1}(0.6)]$
1	秀	$[C^{-1}(0.6), C^{-1}(1)]$

表2. 正規化一致得点 $o(.)$ の瓶境界。これらの境界は、正規化一致得点統計の有用な分類法を得るための検査によって設定された。

これらの瓶境界の選択は、ここでの品質の定義と一致する。つまり、認識率が最も悪いサンプルは「不可」に割り当てられる(この場合、すべてのサンプルが誤って一致する)。品質が「可」のサンプルは境界線上にあるサンプルで、一部のサンプルが誤って一致する。品質が「良」のサンプルは、大部分が正しく一致する。品質が「優」および「秀」のサンプルは、ほとんどすべてが正しく一致する。すなわち、品質が「秀」のサンプルはFNMR(方程式2)がひじょうに高いサンプルで、品質が「不可」のサンプルはFMR(方程式3)が高いサンプルである。

#### 4 指紋画像品質の測定

このセクションでは、指紋画像に関する方程式7および8の $L(.)$ および $I(.)$ の実装について説明する。まず、 $L(.)$ をバイオメトリックサンプル $x_i$ に適用して特徴ベクトル $v_i$ を得る。次に、 $v_i$ をニューラルネットワーク $I(.)$ への入力として使用する。 $L(.)$ は、照合アルゴリズムの情報を伝達するバイオメトリックサンプル $x_i$ の特性と特徴を計算することにより実現される。

##### 4.1 特徴抽出

このセクションでは、指紋の特徴抽出について説明する。ここで提案しているバイオメトリックサンプル品質の定義と測定は、適切な特徴ベクトルが定義され、適切に計算されれば、ほかの生体認証モダリティにも適用できる。

一般に使用されている指紋照合アルゴリズムは、凹凸の明瞭度、特徴点の数と品質の測定基準、および画像のサイズに敏感であることが知られている。ここでは、NISTFingerprint Image Software (NFIS) [7]パッケージ(つまり、方程式7の $L(.)$ の実装)を使って特徴を抽出した。NFISのMINDTCTパッケージは、指紋画像を受け入れ、特徴点を自動的に検出する指紋特徴点検出アルゴリズムを備えている。また、特徴点の品質を評価し、画像品質マップを生成する。NFISは、指紋画像を局地的に分析するため、画像を格子状のブロックに分割する。また、各ブロックの品質を評価するため、複数のマップ(方向マップ、低コントラストマップ、低フローマップ、および高曲率マップ)を計算し、その結果を品質マップに要約する。ブロック内のすべてのピクセルに同じ結果が割り当てられる。NFISのアルゴリズムとソフトウェアパラメータは、500ピクセル/インチ(19.69ピクセル/ミリ)でスキャンされ、256階調のグレーに量子化された画像を処理するために設計および設定されたものであることに注意されたい。MINDTCTパラメータとその品質評価での使用方法については、以下で説明する。

#### 4.1.1 MINDTCT への入力

MINDTCTは、ANSI/NIST形式のファイルを読み込み、ファイル構造からグレースケールの指紋レコードを検索する。このアプリケーションは、ANSI/NISTタイプ4、タイプ13、およびタイプ14の指紋画像レコードを処理できる[8]。

#### 4.1.2 画像品質マップの生成

MINDTCTは、画像内の局地的な領域の品質を測定する。これには、画像内の隆線の指向性フローの判定と、低コントラストの領域、低隆線フローの領域、および高曲率の領域の検出が含まれる。このうち最後の3つの状態は、画像内でも特徴点の検出の信頼性が低い不安定な領域を表しており、これらを組み合わせることで画像内の品質のレベルを表すことができる。これらの特性について、それぞれ以下で説明する。

**4.1.2.1 方向マップ:**このマップの目的は、十分な隆線構造を持つ画像の領域を表すことにある。形がよくはっきり見える隆線は、隆線の末端や分岐点を確実に検出するために不可欠である。指紋を局地的に分析するため、画像は格子状のブロックに分割される。ブロック内のすべてのピクセルに同じ値が割り当てられる。あるブロックから隣のブロックに移るときの値の断絶を最小限に抑えるため、ブロックを囲むウィンドウが定義され、各ウィンドウは隣り合うブロックどうしを覆う。画像内のブロックごとに、それを内包するウィンドウがインクリメンタルに回転され、各方向で離散フーリエ変換(DFT)が行われる。詳しい説明は[7]にある。

**4.1.2.2 低コントラストマップ:**低コントラストマップと呼ばれる画像マップは、コントラストの十分に低いブロックが特定される部分について計算される。このマップにより、指紋から画像の背景が切り離され、指紋の汚れやインクの薄い領域が正確にわかる。画像内のコントラストの低いブロックの内部では、特徴点を検出されない。このソフトウェアは、ブロックを囲むウィンドウ内部のピクセル強度の分布を計算する。分布のうち、高低両側の末端の指定された割合(10%)は外れ値の可能性のあるため切り取られ、残りの分布の幅が測定される。ピクセル強度のしきい値は、実際の指紋画像から抽出されたコントラストの高いブロックと低いブロックの練習用サンプルから経験的に導き出されたものである。ピクセル強度のダイナミックレンジが狭いブロックは、コントラストの低い領域として特定される。

**4.1.2.3 低フローマップ:**低フローマップは、最初に有力な隆線フローを割り当てることができなかったブロックを特定する。低フローの領域で検出された特徴点は信頼できない。

**4.1.2.4 高曲率マップ:**曲率が高い領域で検出された特徴点は信頼できない。指紋のコア領域やデルタ領域では特にこのことがいえる。高曲率マップは、指紋の曲率が高い領域であるブロックを特定するために使われる。

**4.1.2.5 品質マップ:**前述のように、低コントラストマップ、低フローマップ、および高曲率マップは、いずれも品質の低い画像領域を特定するが、それぞれ品質の種類が異なる。これらのマップ内の情報は、1つの全体的なマップに統合され、5つの品質レベル(最高が4、最低が0)を含んでいる。背景は得点0で、得点4は指紋のなかで品質がひじょうによい領域を意味する。特定のブロックに割り当てられる品質は、前述の各マップで特定されたブロックにどれだけ近いかどうかで決定される。ここでは、得点0~4に対して黒、暗い灰色、灰色、明るい灰色、および白を割り当てた品質マップのグレースケール画像を示す。

図3は、品質のよい指紋と、その指紋の品質マップを前述のようにグレースケール化した画像の例である。グレースケールの品質マップ画像は、その大部分が白くなっている。図4は、品質の悪い指紋と、そのグレースケールレベル変換の例である。グレースケールの品質マップ画像は、その大部分が暗い灰色か黒になっているのが特徴的である。2つの指紋の品質マップは、付録Bに掲載した。いくつかの指紋画像の品質マップを主観的にチェックしたところ、MINDTCTによって生成された品質マップは、明瞭な凹凸のある領域や特徴点の検出が信頼できる(またはできない)領域を適切に表現しているという確信が得られた。

それぞれの指紋について、MINDTCTを使ってその品質マップを生成した。品質0のブロックは、背景と見なされる。画像または前景の有効なサイズとして、品質1以上のブロックの総数を計算する。次に、品質が1、2、3、および4の前景ブロックの割合を計算する。これらのブロックを、品質ゾーン1、2、3、および4と呼ぶ。品質ゾーン4の数が多い(すなわち、品質ゾーン1および2の数が少ない)指紋画像ほど望ましい。



図3. 品質がよいと主観的に評価される指紋の例

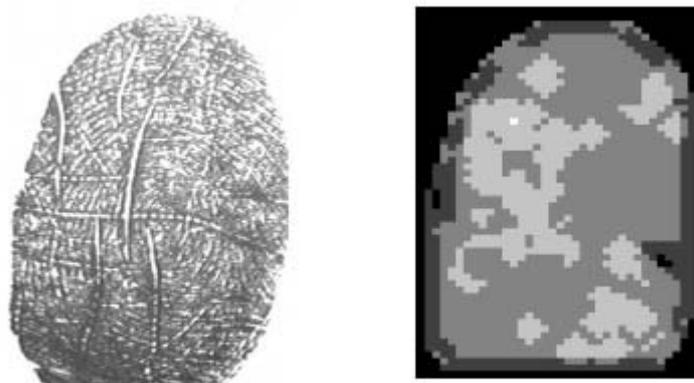


図4. 品質が悪いと主観的に評価される指紋の例

#### 4.1.3 特徴点の品質の評価:

NFISは、検出された個々の特徴点に対応する品質(信頼性)を計算する。NFISは誤った特徴点を削除するためにいくつかの手順を実行するが、通常、誤った特徴点は候補リストに残される。これを管理するには、誤った特徴点に対して正しい特徴点よりも低い品質を割り当てる強固な品質測定基準が有効である。検出された個々の特徴点に対する品質測定基準は、2つの要素を組み合わせで生成される。1つは、前述の品質マップ内の特徴点の場所から直接取得される。もう1つは、特徴点の周辺を範囲とするサンプルのピクセル強度統計値(平均および標準偏差)に基づいている。明瞭

な凹凸がある領域は、完全なグレースケールのスペクトルをカバーする高いコントラストを持つ。その結果、周辺部の平均ピクセル強度は127にきわめて近くなる。同じ理由で、理想的な周辺部のピクセル強度は標準偏差が64以上になる。NFISは、この理屈に基づき、4.1.2.5で説明した品質マップを使用して、各特徴点に0.01～0.99の範囲の品質値を割り当てる。低い品質値は特徴点が品質の低い画像領域で検出されたことを示し、高い品質値は特徴点が品質の高い画像領域で検出されたことを示す。

1つの指紋について、NFISは個々の特徴点の品質を検出し、評価する。品質が0.5より低い特徴点は信頼できない。ここでは、品質が0.5以上、0.6以上、0.75以上、0.8以上、および0.9以上の特徴点の数を計算する。

#### 4.1.4 特徴ベクトル

NFISの品質マップと特徴点の品質評価を使用して、表3に示す特徴ベクトルを定義した。このため、NFISのMINDTCTを使って、個々の指紋について表3に示す11次元の特徴ベクトルを計算する。

	名前	説明
1	前景	品質が1以上のブロックの数。つまり、前景 = $\sum_{i=1}^4 U_i$ ( $U_i$ は品質が <i>i</i> のブロックの数)
2	特徴点の総数	指紋内に見つかった特徴点の総数
3	特徴点 05	品質が0.5以上の特徴点の数
4	特徴点 06	品質が0.6以上の特徴点の数
5	特徴点 075	品質が0.75以上の特徴点の数
6	特徴点 08	品質が0.8以上の特徴点の数
7	特徴点 09	品質が0.9以上の特徴点の数
8	質ゾーン 1	品質が1である品質マップの前景ブロックの割合
9	品質ゾーン 2	品質が2である品質マップの前景ブロックの割合
10	品質ゾーン 3	品質が3である品質マップの前景ブロックの割合
11	品質ゾーン 4	品質が4である品質マップの前景ブロックの割合

表3. 特徴ベクトルの説明

図3の(品質がよいと主観的に評価される)指紋の特徴ベクトルは、次のように計算される。

00602\_01: (1642, 57, 34, 34, 34, 34, 28, 0.053, 0.167, 0.059, 0.720)

また、図4の(品質が悪いと主観的に評価される)指紋の場合は、次のようになる。

00207\_01: (1690, 182, 0, 0, 0, 0, 0, 0.168, 0.572, 0.259, 0.0006)

明らかに、品質のよい指紋の場合は品質ゾーン4が大きく、信頼性が0.9以上の特徴点の数が多いのに対し、品質の悪い指紋の場合は品質ゾーン2が大きく、信頼性が0.5以上の特徴点が1つもない。

特徴ベクトルの本質を理解し、その基本構造を明らかにするため、探索的データ分析(EDA)を行った。EDAは、古典的な統計的推測の方法で必要とされる仮説(外れ値がなく、ほぼ正規分布の、系列相関でない観察結果)がデータによって裏付けられるかどうかを確認するための、グラフィック指向のデータ分析方法である。ここでは、データセットDOS-Cの右手人差し指について計算された特

徴ごとに、ヒストグラム、ボックスプロット、密度プロット、正規QQプロットの4つのグラフを示す。正規QQプロット(四分位-四分位プロット)は、値の順に並べたデータと、それに対応する標準正規分布の分位から成る。QQプロットがほぼ直線になる場合、データは適度にガウス分布である。直線にならない場合、データはガウス分布でない。4つのプロットのうち、ヒストグラムと密度プロットでは分布の形状が最もわかりやすく示され、ボックスプロットと正規QQプロットでは外れ値がはっきりと示される。各ボックスプロット内のボックスはデータの第1分位と第3分位を表し、ボックス内の黒い点は中央値を表し、第1分位と第3分位の範囲外にあるデータ点はボックスの外側の黒い線で表示される。図5に、データセットDOS-C内の右手人差し指の指紋に関する分析を統計図によって要約したものを示す。表4に、データセットDOS-Cの右手人差し指の指紋の特徴に関するスピアマン順位相関と、方程式4で定義されたベンダーFの正規化一致得点を示す。

	品質 ゾーン1	品質 ゾーン2	品質 ゾーン3	品質 ゾーン4	正規化 一致 得点
前 景	0.139	0.139	0.012	0.079	0.108
特徴点 の総数	0.258	-0.385	0.292	-0.315	0.076
特徴点 05	-0.347	-0.386	-0.510	0.563	0.391
特徴点 06	-0.349	-0.387	-0.513	0.566	0.391
特徴点 075	-0.412	-0.253	-0.521	0.555	0.313
特徴点 08	-0.398	-0.169	-0.455	0.483	0.246
特徴点 09	-0.339	-0.047	-0.303	0.327	0.131
正規化 一致 得点	-0.242	-0.172	-0.255	0.289	1.0

表4. 特徴と正規化一致得点の相関—データセットDOS-C—右手人差し指の指紋とベンダーFの得点。特徴点06と品質ゾーン4の相関が最も高い(0.566)。正規化一致得点と特徴のあいだに有意な相関はない。

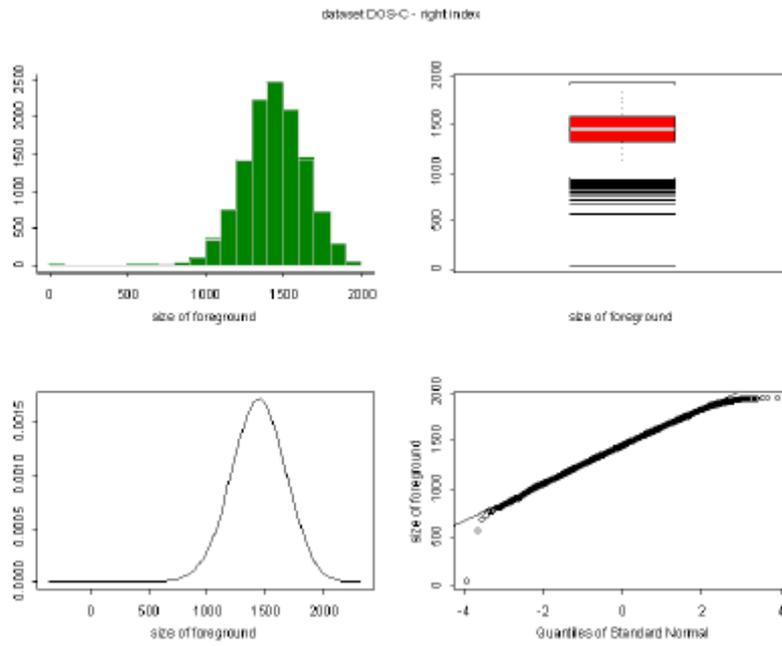


図5-a. 前景(品質が1以上のブロックの数)に関する探索的データ分析

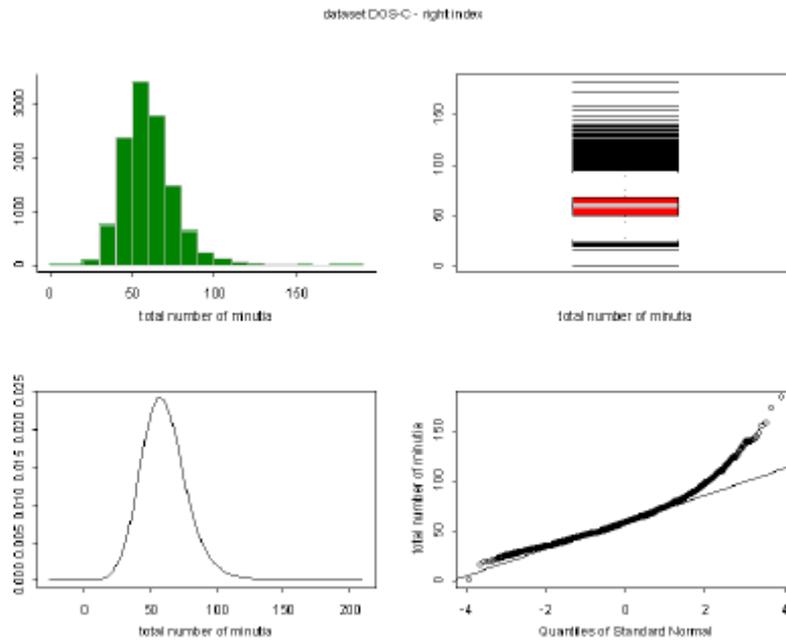


図5-b. 特徴点の総数に関する探索的データ分析

データセットDOS-Cの右手人差し指

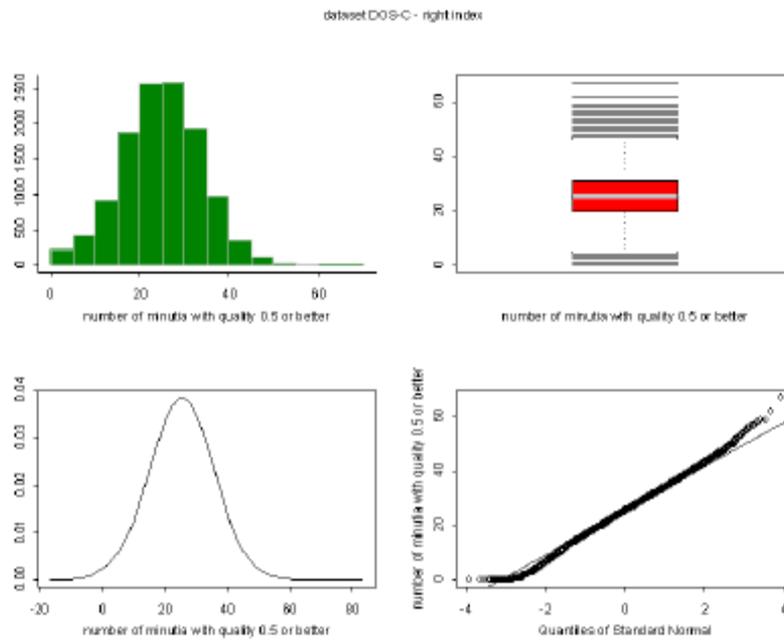


図5-c. 特徴点05(品質が0.5以上の特徴点の数)に関する探索的データ分析

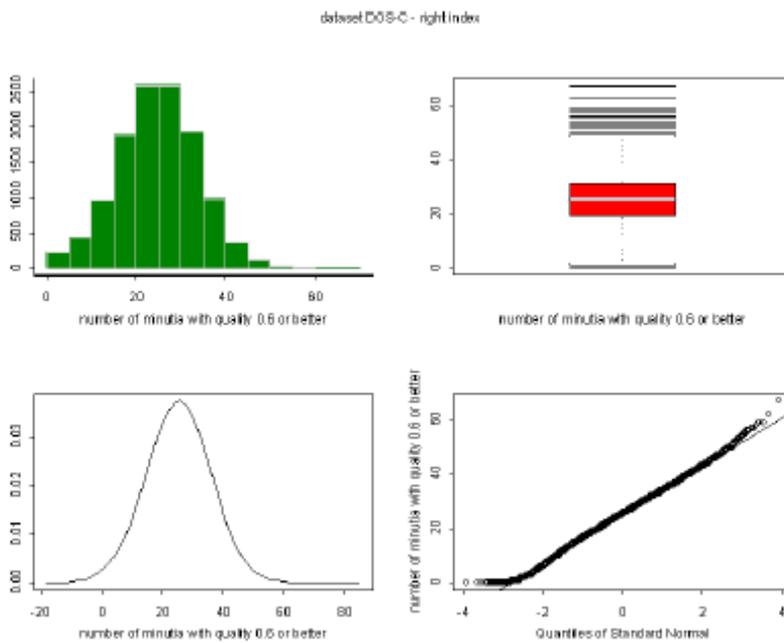


図5-d. 特徴点06(品質が0.6以上の特徴点の数)に関する探索的データ分析

データセットDOS-Cの右手人差し指

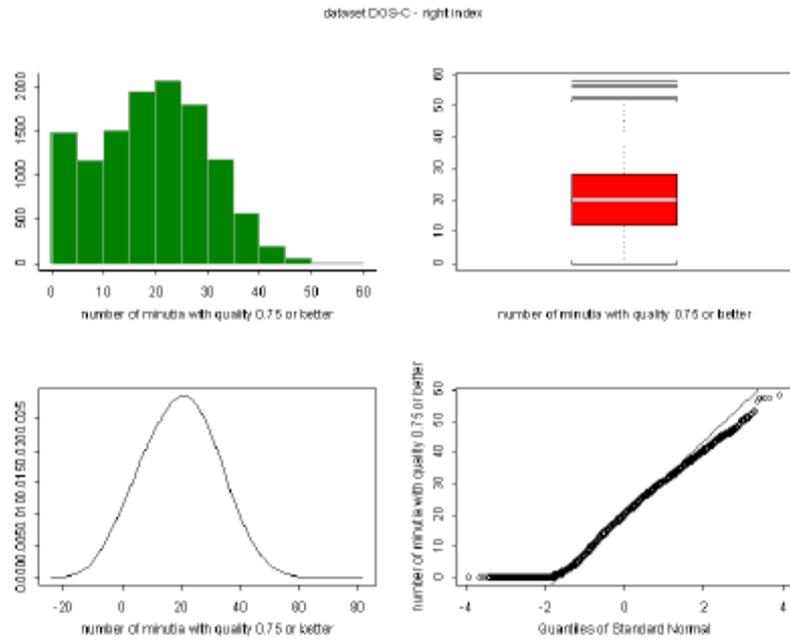


図5-e. 特徴点075(品質が0.75以上の特徴点の数)に関する探索的データ分析

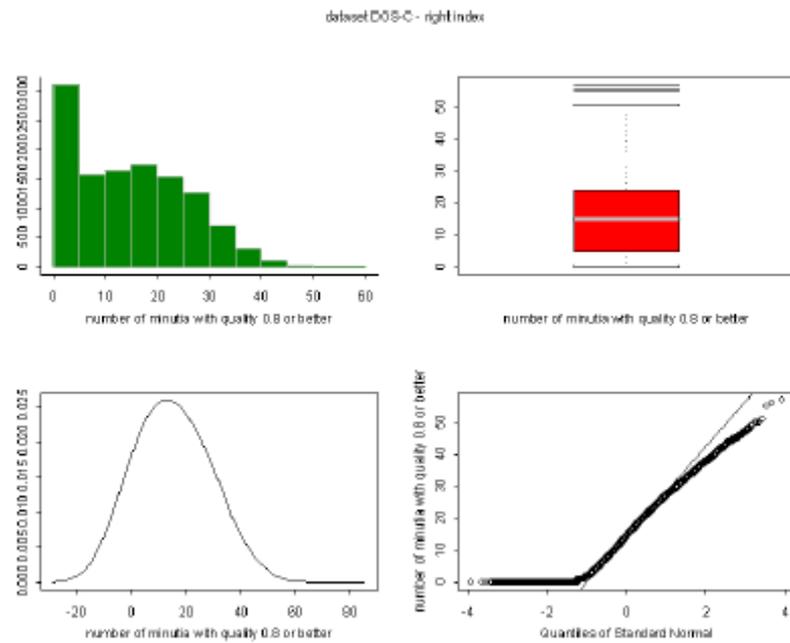


図5-f. 特徴点08(品質が0.8以上の特徴点の数)に関する探索的データ分析

データセットDOS-Cの右手人差し指

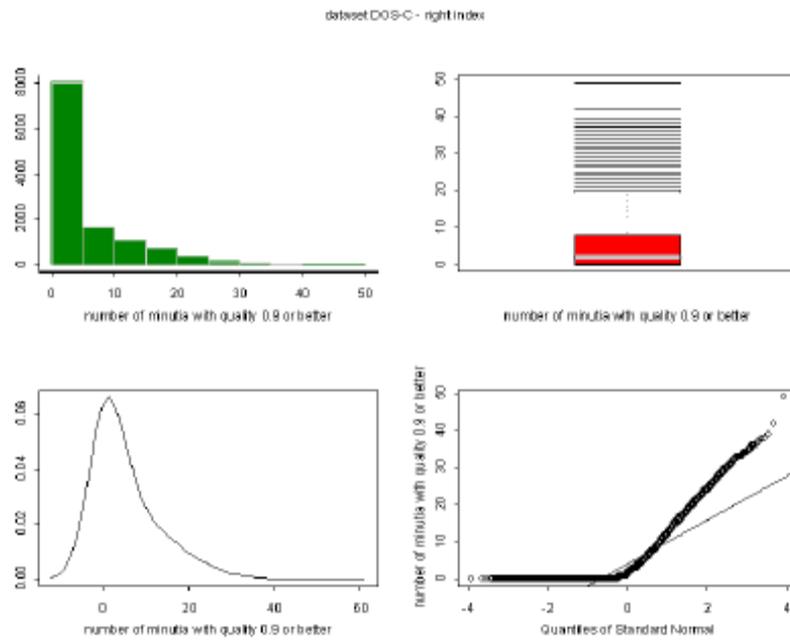


図5-g. 特徴点09(品質が0.9以上の特徴点の数)に関する探索的データ分析

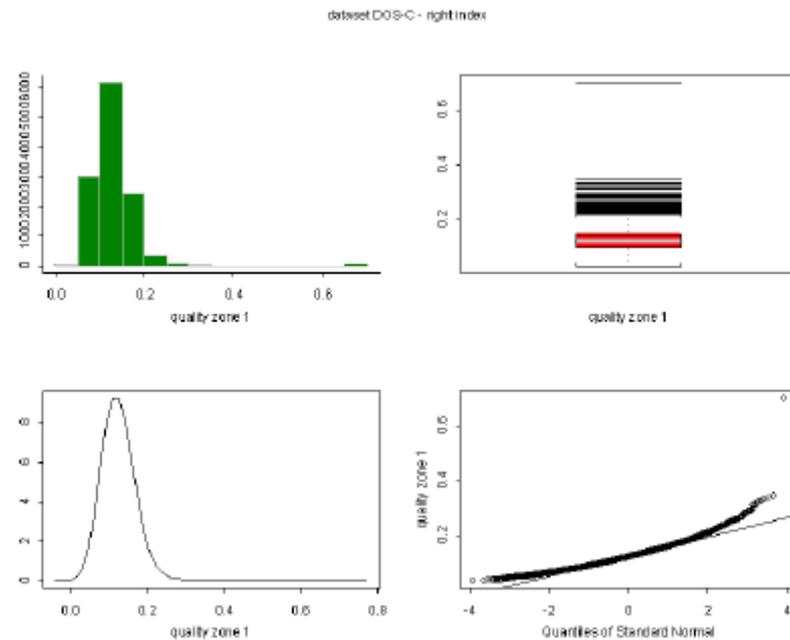


図5-h. 品質ゾーン1(前景のうち品質が1の部分)に関する探索的データ分析

データセットDOS-Cの右手人差し指

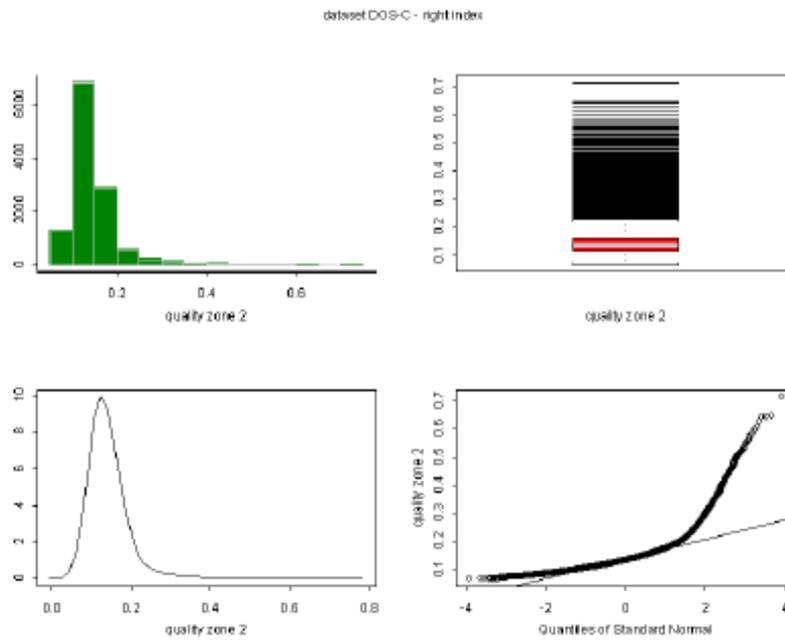


図5-i. 品質ゾーン2(前景のうち品質が2の部分)に関する探索的データ分析

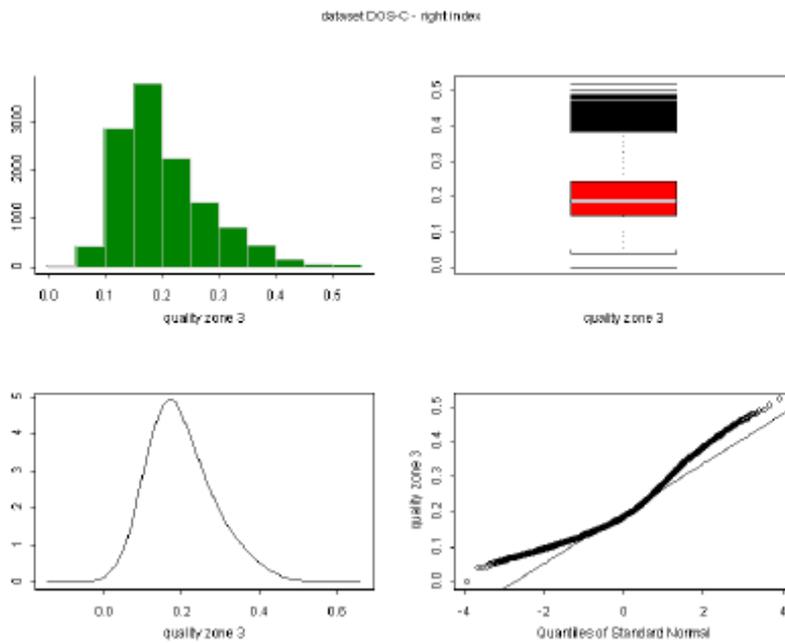


図5-j. 品質ゾーン3(前景のうち品質が3の部分)に関する探索的データ分析

データセットDOS-Cの右手人差し指

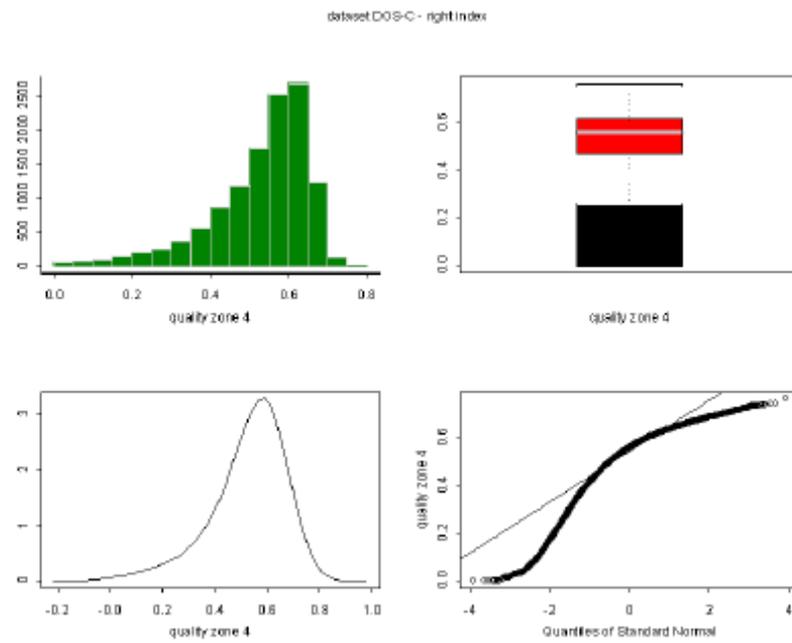


図5-k. 品質ゾーン4(前景のうち品質が4の部分)に関する探索的データ分析

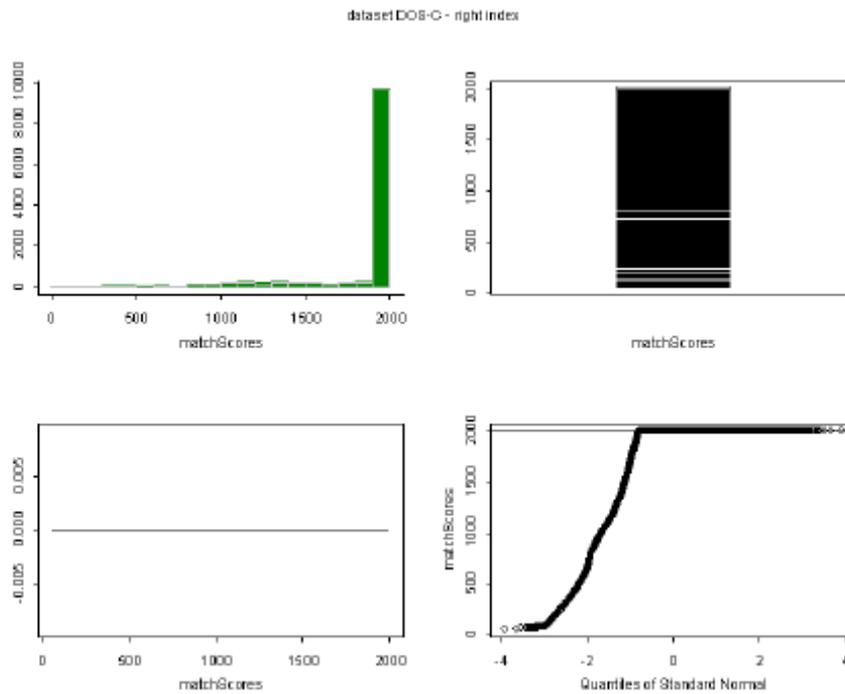


図5-l. ペンダーFの一致得点に関する探索的データ分析  
データセットDOS-Cの右手人差し指、ペンダーFの正規化一致得点

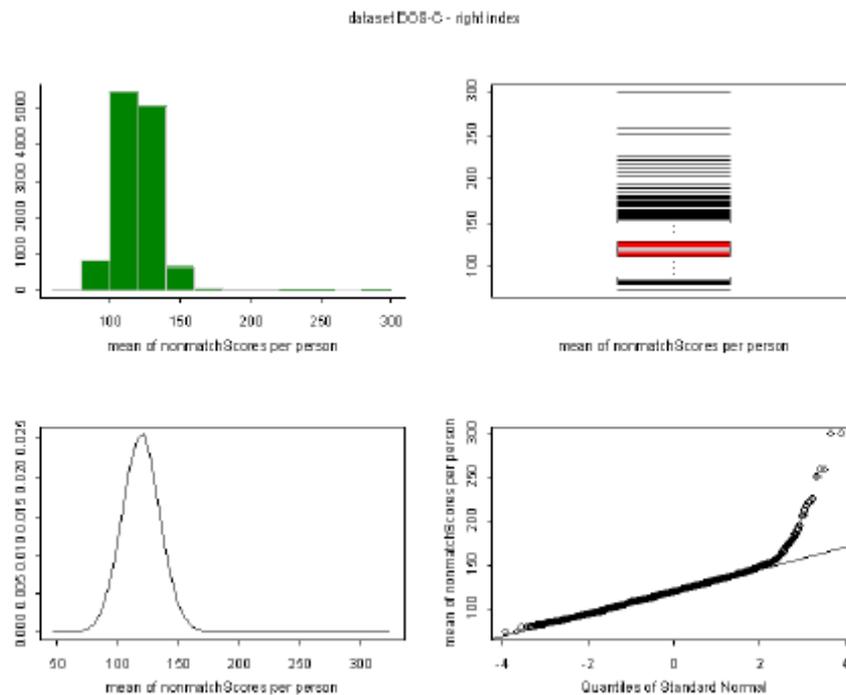


図5-m. ベンダーFの不一致得点に関する探索的データ分析ーデータセットDOS-C右手人差し指

不一致得点の1人当たりの平均の末端が正規分布からはずれていることに注意されたい。前述のように、データセットDOS-Cは統合されているため、不一致の平均点は誤りを含む実観測データによって人為的に増加していない。

図5から、次のようにいくつかの興味深い観察結果がわかる。

○ 外れ値はデータに大きな影響を及ぼす。 図5-a～5-mに示したように、外れ値の存在は強固な統計的手法が必要であることを示している。観察結果の断定的な取捨選択には無駄が多いため、データを選別して外れ値を取り除くだけでは十分でないことに注意する必要がある。また、多変量データに含まれる外れ値を見きわめるのは、ひじょうに困難か、不可能である。外れ値を取り除くと、分布にも影響する。誤った観察結果は完全に破棄したくなるものだが、信頼できない観察結果を破棄せずに、その比率をさげることで、より適切に対応できる。強固な測定基準を算出するには、刈り込み平均または順位相関を計算する。刈り込み平均は分布の中央部 $(1-2\alpha)$ の平均であるため、各末端から観察結果の断片 $\alpha$ が取り除かれる。

○ 分布が正規分布でない。 また、正規分布に近くもない。このことから、ノンパラメトリック法を使用すべきであることがわかる。

図5-a～mのQQプロットからわかるように、正規分布は仮定として妥当でないため、表4ではDOS-Cの右手人差し指サンプルの特徴とベンダーFの正規化一致得点(つまり $o(\cdot)$ )のあいだの相関について、順位ベースのスピアマンの $\rho$ 測定基準を計算し、報告している。

特徴と正規化一致得点のあいだの相関係数は、あまり興味を引くものではない。特徴点05(0.391)、特徴点06(0.391)、および品質ゾーン4(0.289)が最高である。これは、どの特徴もそれ自体では正規化一致得点の適切な予測材料にならないことを意味している。しかし、特徴を組み合わせることによって予測できる可能性はある。特徴と正規化一致得点のあいだに何らかの関連(相関)があるかどうかを調べるため、個々の特徴と正規化一致得点の散布図を作成した。散布図は、2つの変数間の相関を視覚的にチェックする手段として使用できる。2つの変数が強く相関している場合、一方の変数に対するもう一方の変数の散布図は直線状になる。ある特徴と正規化一致得点( $o(\cdot)$ )のあいだに強い相関があるということは、その特徴を正規化一致得点 $o(\cdot)$ の予測材料として使用できることを意味する。図6、7、8、および9に、DOS-Cの右手人差し指の指紋に関する特徴ベクトルの成分とベンダーFの正規化一致得点の散布図の行列を示す。緑の丸は正しく一致する指紋を表し、赤い丸は正しく一致しない指紋を表す。

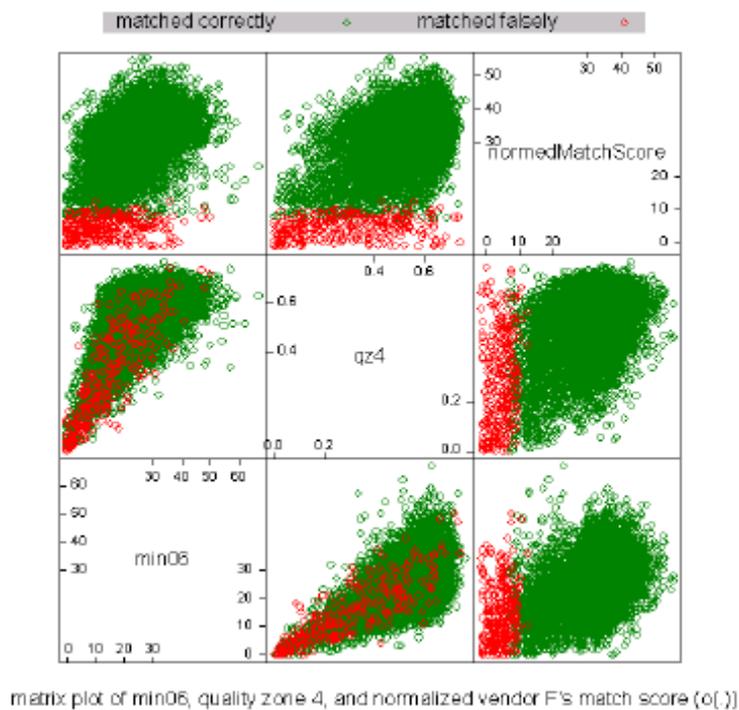
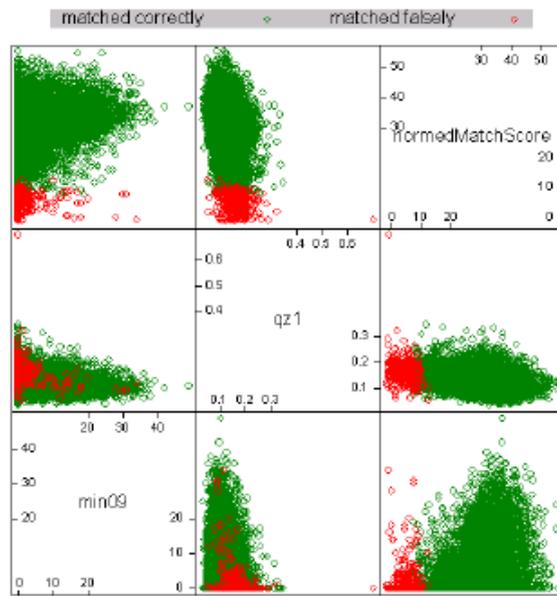
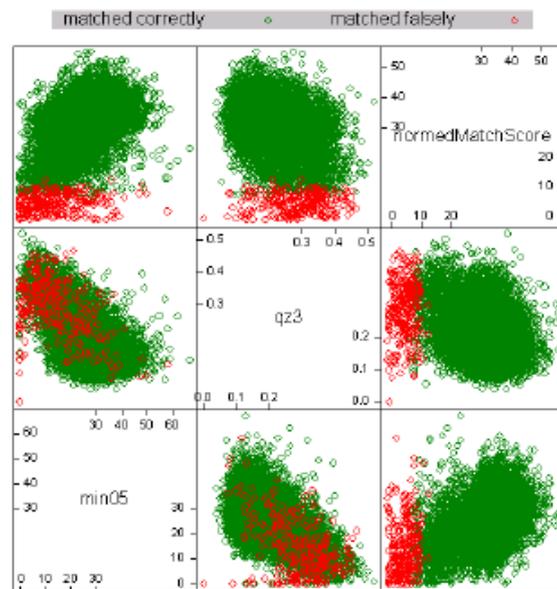


図6. 特徴点06と品質ゾーン4には何らかの相関関係がある。特徴点06と品質ゾーン4の値が小さい指紋は、誤って識別される可能性が最も高い。



matrix plot of min09, quality zone 1, and normalized vendor F's match score (o[.])

図7. DOS-Cの右手人差し指に関する特徴点09、品質ゾーン1、およびベンダーFの正規化一致得点の散布図。特徴点09が15を超える指紋は、正しく識別される可能性が最も高い(したがって、われわれの定義では、品質が「よい」)。



matrix plot of min05, quality zone 3, and normalized vendor F's match score (o[.])

図8. 特徴点05の値が小さく、品質ゾーン3の値が大きい指紋は、誤って識別される可能性が最も高い。

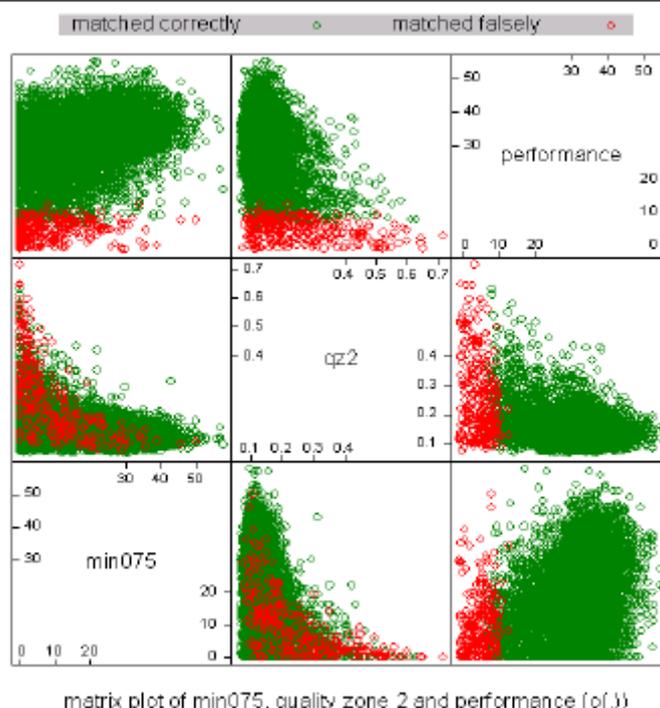


図9. 特徴点075の値が大きく(20を超え)、品質ゾーン2の値が小さい指紋は、正しく識別される可能性が最も高い(したがって、品質が「よい」)。

散布図は、表4で説明した特徴相関の結果をはっきりと示している。図6は、方程式4を使って特徴点06、品質ゾーン4、およびベンダーFの正規化一致得点のあいだで生成された3種類の散布図を示している。この図では、特徴点06または品質ゾーン4と正規化一致得点とのあいだに明確な関係はみられない。しかし、2つの弱い相関が示されている。図6の右下の散布図では、特徴点06の値が十分に大きいときは緑の丸になっており、正しい一致が予測される。さらに、図6の左中央と中央下の散布図では、品質ゾーン4の値と特徴点06の値がいずれも小さいときは誤った一致(赤い丸)になっている。このような性質は、特徴点075と品質ゾーン2の特徴に関する図9でもみられる。図8は、特徴点05と品質ゾーン3のあいだに弱い相関があることを示している。図8の中央下にある散布図の赤い丸は、品質ゾーン3の値が十分に大きく、特徴点05の値が小さい場合は誤って識別されることを示している。特徴点09と品質ゾーン1の特徴に関する図7は、特徴点変数の値が大きい場合の相関を示しているが、特徴点変数と品質変数が組み合わされた場合の予測効果は示していない。正規化一致得点に対するほかの特徴の効果についても、同じ結果が出ている。

NFISによって生成された品質測定基準が(使われている特徴抽出方法がまったく異なる)商用の指紋照合システムの性能を予測できるという観察結果は、きわめて重要である。これは、NFISによって生成され、一般に公開されている特徴に基づく画像品質の測定基準を使って、1つの(また、セクション5で示すように、ほかの)商用システムの照合性能を予測できることを示している。

要するに、単独で正規化一致得点を予測するのに必要な相関係数が得られるような特徴は存在しない。しかし、これらの特徴の線形または非線形の結合により、方程式4で定義した正規化一致得点を予測できるという強力な証拠が存在する。次のセクションでは、これらの特徴を使って正規化一致得点を予測する方法について説明する。

## 4.2 パターン認識(回帰と分類)

パターン認識の用語では、既知の例が(それらの分類とともに)トレーニングセットと呼ばれ、将来のケースがテストセットになる。成功の第一の尺度は、誤差(または誤分類)率である。トレーニングセットを再分類すると(おそらくかなり)偏った推定値が得られるが、母集団全体からランダムに選んだテストの誤差率は不偏推定量になる。過学習の問題を避けるため、データセットDOS-C、DHS2-C、BEN、TXDPS、DHS10から取り出した指紋画像の互いに素なサブセットを2つ作成し、一方のサブセット(トレーニングセット)をモデルを当てはめるために使用し、もう一方のサブセット(テストセット)を最終的に選ばれたモデルの汎化誤差を評価するために使用した。検証セットとして、データセットSD29[3]を使用した。

方程式8は、各種の回帰法を使用すべきであることを示している。ここでは、ベクトル入力 $v$ =(特徴 $_1$ , 特徴 $_2$ , ..., 特徴 $_n$ )があり、その実数値出力 $o(\cdot)$ を予測したい。線形回帰モデルは次のように表される。

$$Y = f(x) = \beta_0 + \sum_j \beta_j x_j$$

この線型モデルは、回帰関数 $E[Y|X]$ が線形であること、またはこの線型モデルが妥当な近似であることを仮定している。 $\beta_j$ は未知のパラメータまたは係数であり、変数 $x_j$ はさまざまなソースから得られる。これらのソースには、定量的入力、定量的入力の変換(対数、平方など)、基底展開( $x_j$ など)、および変数間の相互作用( $x_i x_j$ )がある。 $x_j$ のソースに関係なく、このモデルはパラメータ $\beta_j$ に対して線形である。

グラフィックデータによって応答(正規化一致得点)と予測材料(特徴)の直線関係がはっきりとわかる場合、線形回帰は自然な選択である。データセットDOS-Cの散布図(図5-a~m)や図6、7、8、9の行列散布図によれば、このような線形関係は存在しない。いくつかの回帰法を試みたが、データに当てはまるものはなかった。 $R^2$ が0.30を超えることはなく、しかも残差に標準の共分散構造がなく、残差の広がり当てはめ値よりも大きかった。このため、伝統的な線形モデルは(簡単ではあるが)われわれのデータに対する適切な手法ではない。これは、伝統的な線形回帰の手法では基礎データに関していくつかの強い仮定が立てられるが、特徴データはいくつかの点でこれらの仮定を満たさないためである。たとえば、1つ以上の外れ値が回帰直線から外れている場合や、データが正規分布にならない場合がある。外れ値は、データに大きな影響を及ぼし、正規分布にならないことがわかっている。外れ値の影響を克服するため、ここでは、線形法の基本形式を維持しながら、外れ値の影響を最小限に抑える、強固な回帰法を使用した。さらに、1つ以上の予測材料を予測材料に関する多項式の滑らかな関数によって置き換えることにより、伝統的な線型モデルを一般化加法モデル(GAM)に変換した。回帰の設定では、一般化加法モデルが次のように表される[23]。

$$Y = \alpha + f_1(x_1) + f_2(x_2) + \dots + f_p(x_p) + \epsilon$$

例によって、 $x_1, x_2, \dots, x_p$ は予測材料(特徴)を表し、 $Y$ は結果(正規化一致得点)である。 $f_i$ は特定されていない滑らかな(したがってノンパラメトリックな)関数である。ここで採用したアプローチは、散布図の平滑化(LOWESS平滑化やスプライン平滑化など)を使って各関数を当てはめ、すべての $p$ 関数を同時に求めるためのアルゴリズムを提供する方法である。このモデルは、バックフィッティングによって加重加法モデルを繰り返し当てはめる局所的な得点アルゴリズムを使って当てはめられた。バックフィッティングアルゴリズムは、部分的な残差を繰り返し平滑化することによって加法モデルを当てはめる、一種のGauss-Seidel法である。このアルゴリズムは、フィットのノンパラメトリックな部分からパラメータ部分を分離し、バックフィッティングアルゴリズムの内部で重み付き線形最小二乗法を使ってパラメータ部分を当てはめる。このモデルに対してLOESS平滑化とスプライン平滑化を試

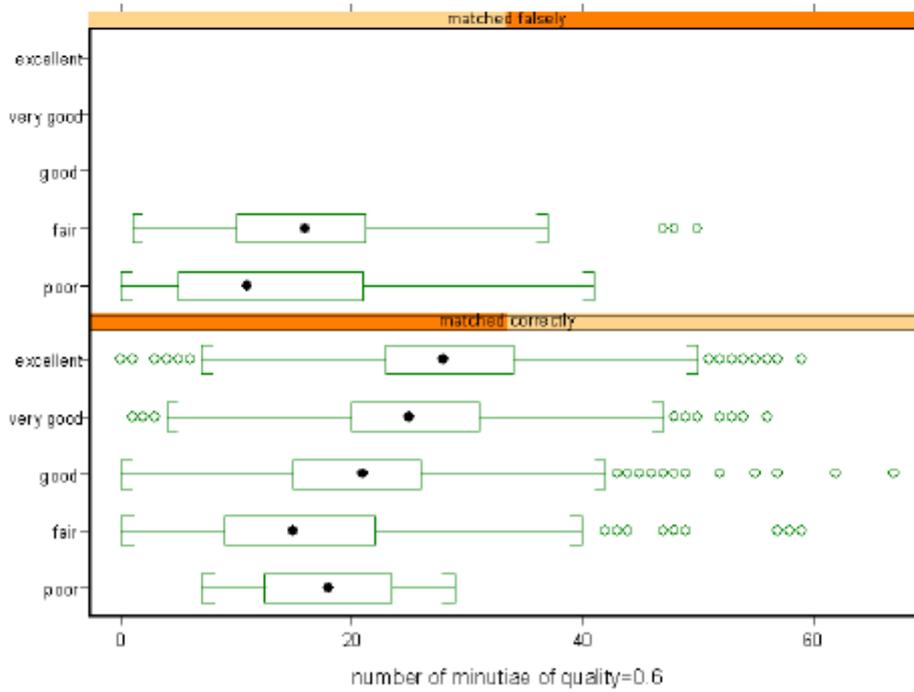
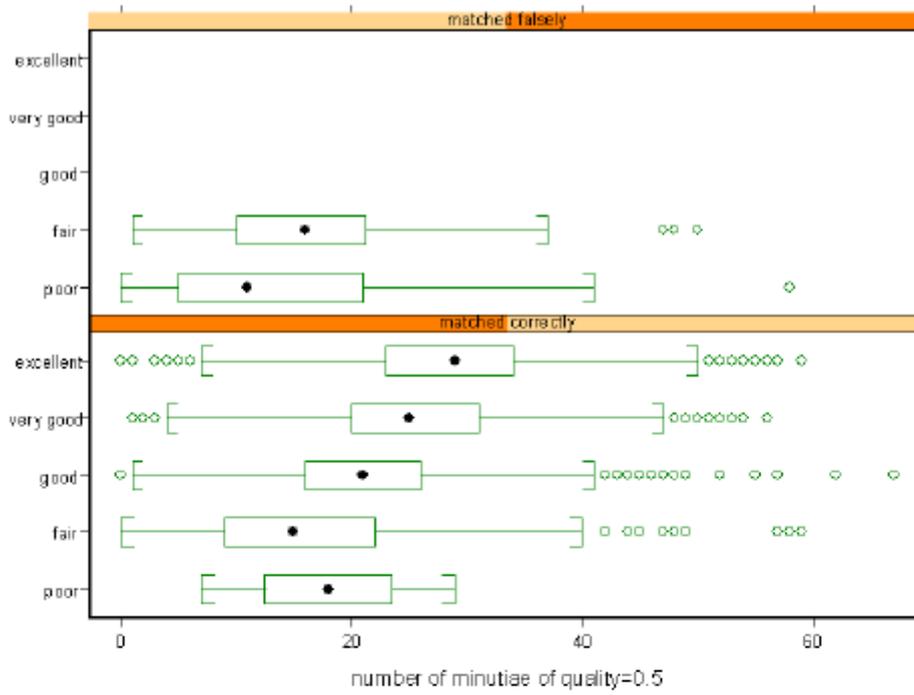
みた。GAMIは誤差(残差)の非線形性や非正規性に対応するが、予測材料 $x_j$ (つまり特徴)間の相互作用のために失敗した。このことは、次に射影追跡回帰を使用すべきであることを示している[4]。これは、説明変数間に相互作用がある可能性を含むノンパラメトリック回帰モデルである。射影追跡回帰は、射影された変数に加法モデルを適用する。つまり、次のように表される。

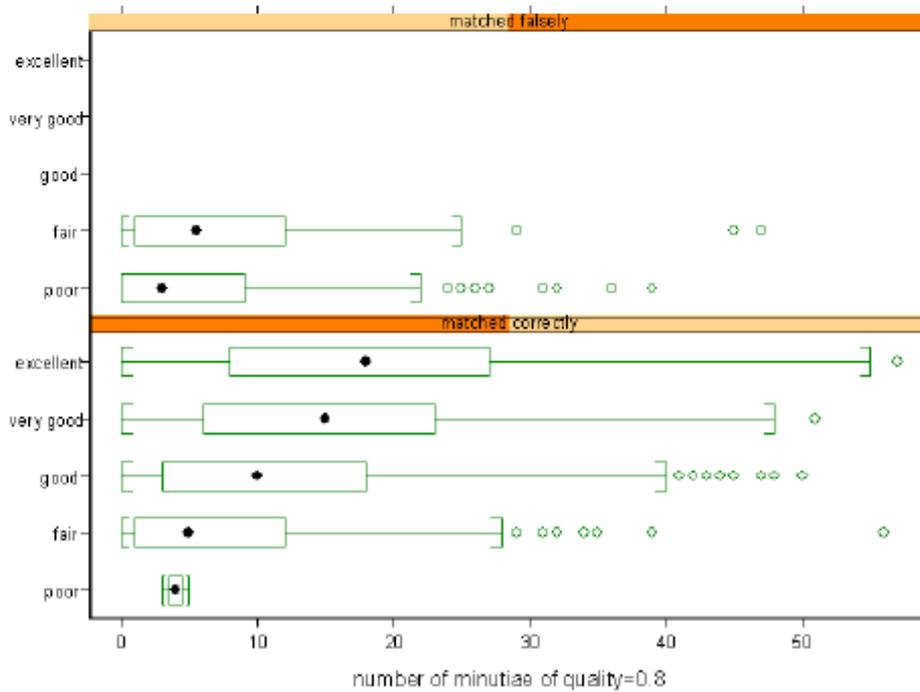
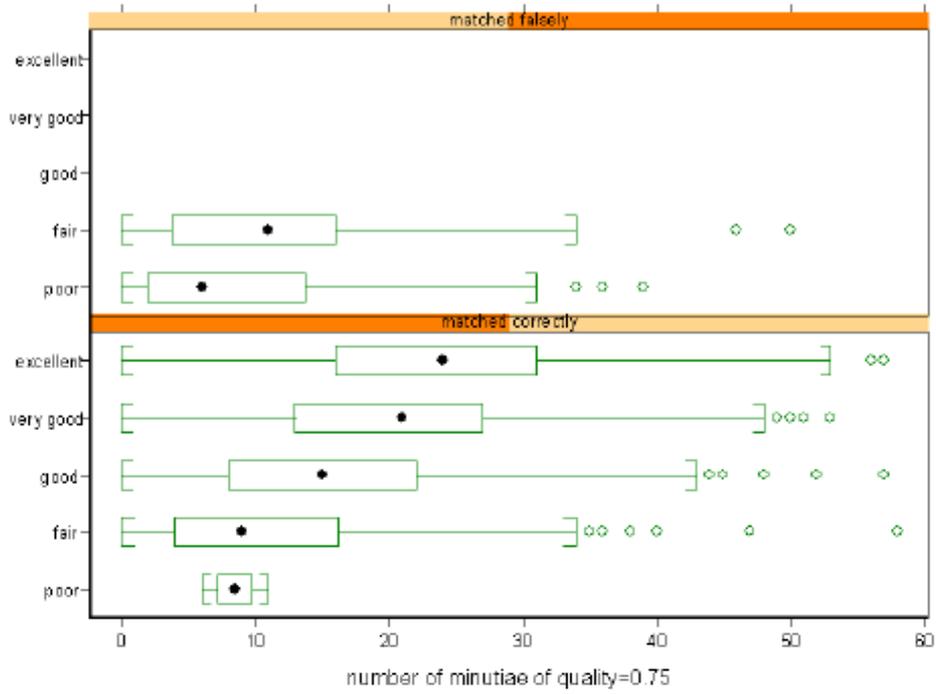
$$Y = \alpha_0 + \sum_{j=1}^M f_j(\alpha_j^T X) + \varepsilon$$

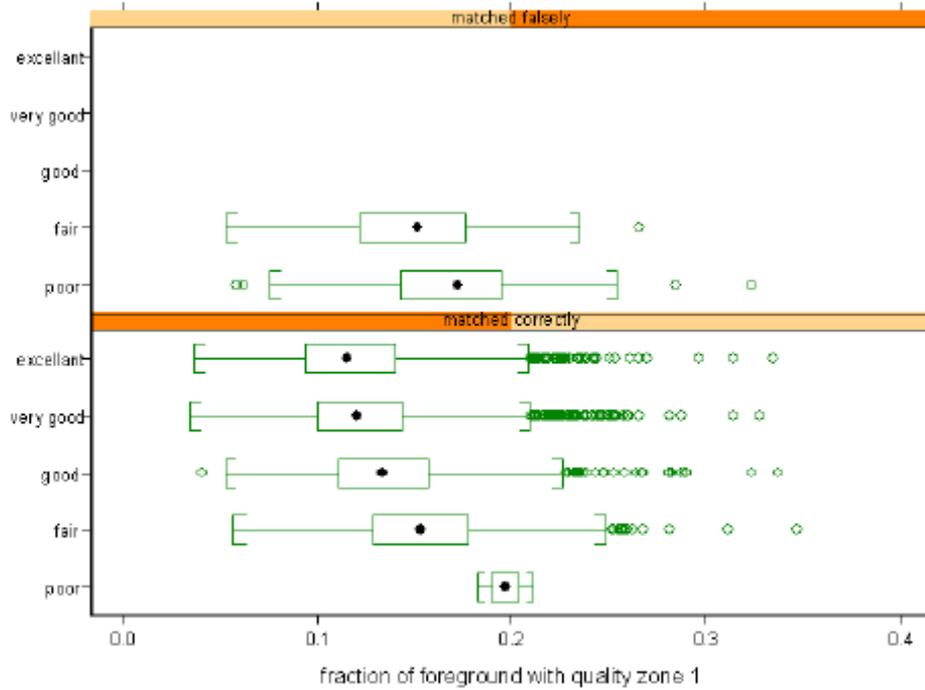
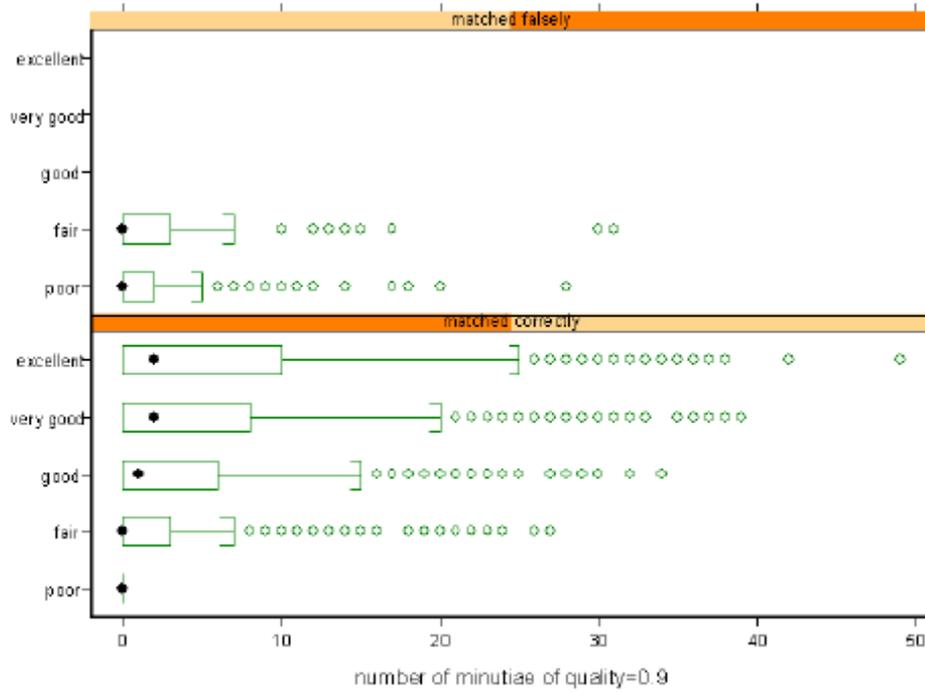
ここで、ベクトル $\alpha_j$ と次元 $M$ はユーザによって選択される。このように、 $M$ の $X$ を注意深く選んだ方向に射影することによって形成された予測材料変数に対して、加法モデルが使用される。射影追跡回帰でも、これまでに示したほかのモデルより適切な結果は得られなかった。

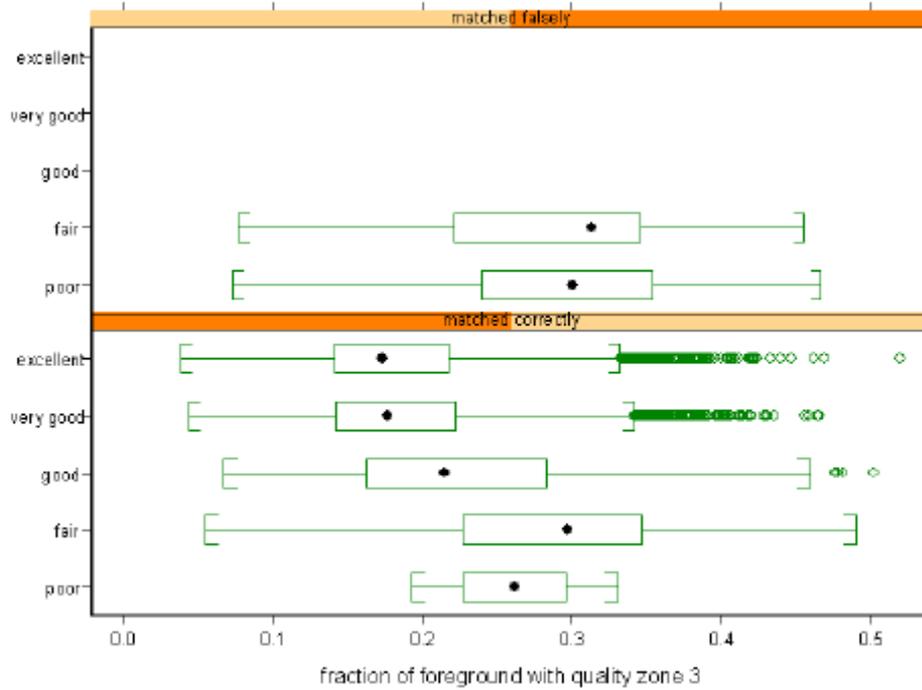
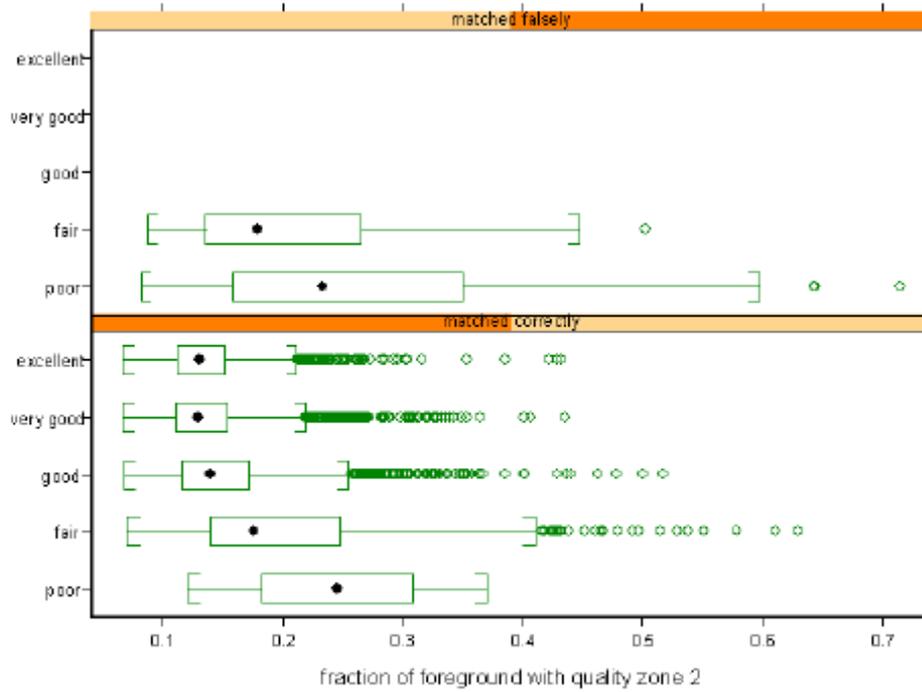
回帰モデルや加法モデルの失敗に加え、離散的な品質を選択する論理的根拠(セクション3.1)も存在するため、この問題を分類の問題として提示することにする。特徴ベクトルの識別力を調べるため、データセットDOS-Cの右手人差し指の指紋について、ベンダーFの正規化一致得点のレベル(表2に示した瓶)に応じて、正しく識別されたものと誤って識別されたものを別々に示したボックスプロットを特徴ごとに作成した。これらのボックスプロットを図10に示す。各ボックスプロット内のボックスはデータの間を示し、ボックス内部の黒い点は中央値を表す。ひげは1.5(四分位範囲)にあり、範囲外の点は外れ値と見なされ、個別に描画されている。ボックスプロットは、データの場所や広がりを示すだけでなく、ひずみも示している。データセットDOS-Cの右手人差し指の指紋に関する特徴ベクトルの成分とベンダーFの正規化一致得点のトレリス図を付録Bに掲載した。

図10に示すように、各特徴の中央値は、性能の瓶ごとに異なる。このことは、これらの特徴を使って、方程式4で定義した正規化一致得点の空間を分割できることを暗に示している。









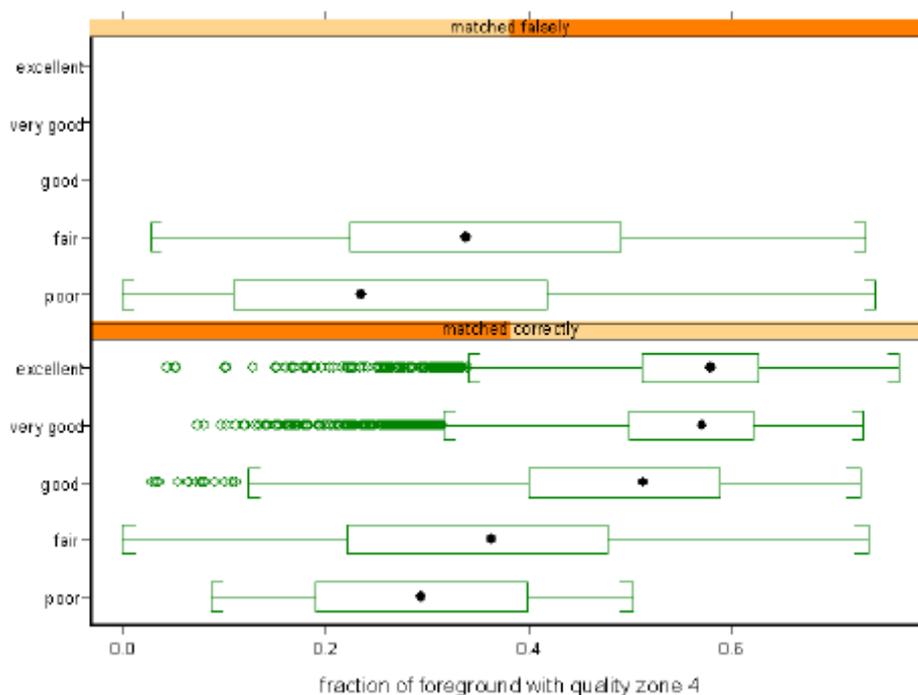


図10. データセットDOS-Cの右手人差し指の指紋に関する9つの特徴を、それぞれベンダーFの正規化一致得点の瓶ごとに示したボックスプロット。各特徴の中央値は、得点の瓶ごとに異なる。「良」、「優」、「秀」の瓶に含まれるサンプルは、正しく一致する。

ここでは、非線形の分類方法として、人工ニューラルネットワークを選択した。ニューラルネットワークは、任意の非線形関数に対する近似関数の役割を果たすことができ、特徴や分類空間に含まれるモデルベースの分布関数に依存しない。

#### 4.2.1 ニューラルネットワークのトレーニング

ここでは、NFISパッケージに含まれるニューラルネットワークの実装を使用した。このプログラムで使われている機械学習法の背後にある理論については、[17]で説明されている。このプログラムは、3層フィードフォワード非線形パーセプトロンモデルをトレーニングする。入力ノードは隠れノードに、隠れノードは出力ノードに、それぞれ正方向送りで接続される。重みのトレーニング(最適化)は、スケールされた共役勾配法アルゴリズム[18]を使用するか、または最初にスケールされた共役勾配法を使用し、次に省メモリのBFGS(Broyden Fletcher Goldfarb Shanno)法アルゴリズム[19]に切り替えることによって行われる。トレーニング中は、ボルツマンの枝刈り[20](つまり、接続の削除)が行われる。

セクション3.1(表2)では、5種類の品質値について説明した。これら5つの品質値は、ニューラルネットワークの出力である。ニューラルネットワークへの入力、セクション4.1.4で説明した11次元の特徴ベクトルである隠れノードは22個選択した。隠れノードと出力ノードに使用される活性化関数は、正弦関数である。また、ボルツマンの枝刈りを設定した。

トレーニングセットは、データセットDOS-C、DHS2-C、DHS10、TXDPS、およびBENの右手人差し指、左手人差し指、右手親指、および左手親指の平面押捺指紋で構成される。セクション2.1で説明したように、これらのデータセットには、異なる担当者が異なる場所や条件下で収集した指紋が含まれている。そのため、トレーニングセットにはさまざまなレベルの品質が含まれている。ここでは、VISIT\_POEデータセット、VISIT\_POE\_BVAデータセット、およびDOS-C、DHS2-C、DHS10、TXDPS、およびBENのうちトレーニングに使用しなかったサブセットについて、システムをテストした。

ニューラルネットワークのトレーニングには、トレーニングセットの完全な類似度行列が必要である。ニューラルネットワークの出カクラス(つまり、正規化一致得点の瓶)を計算するには、類似度得点が必要である。瓶境界は一致得点と不一致得点の分布に基づいているため(表2)、1つのベンダーの類似度得点に基づいてトレーニングを行うと「ベンダー依存」の重みが付き、システム全体に偏りが生じるため、ほかの照合プログラムにシステムを一般化できないのではないかという議論もある。このような問題を避けるため、SDKベンダーF、G、およびVTBが同じ瓶に割り当てたサンプルを選んだ。したがって、トレーニングにはSDKベンダーF、G、およびVTBの類似度得点を使用した。

異なるデータセット、品質、および指位置の間で数のバランスが取れたトレーニングセットを設計するように心がけた。表5は、トレーニングセットの散らばり具合をまとめたものである。

品質	右手人差し指					左手人差し指					右手親指					左手親指				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
DOS-C	100	100	83	21	31	100	100	69	18	62										
DHS2-C	100	100	90	29	21	100	100	79	11	30										
BEN	40	40	40	40	26	40	40	40	40	40	60	60	60	60	7	60	60	60	60	15
TXDPS	40	40	40	38	35	40	40	40	36	34	60	60	60	43	14	60	60	60	31	18
DHS10	20	20	20	20	20	20	20	20	20	20	80	80	80	32	32	80	80	80	36	39
合計	300	300	273	148	133	300	300	248	125	186	200	200	200	135	53	200	200	200	127	72

図5. トレーニングセットの構造: サンプルの数。  
クラスごとの合計は、順に1000、1000、421、535、および444である。

理想的には、トレーニングセットに含まれる指(右手人差し指、左手人差し指、右手親指、左手親指)ごとの数を同じにし、品質ごとの数も同じにしたい(たとえば、右手人差し指について品質1、2、3、4、5の指紋をそれぞれ300個とし、左手人差し指、右手親指、左手親指についても同様とするなど)。さらに、前述のデータセットごとのサンプル数も同じになるのが望ましい。しかし、現実には品質のきわめて悪いデータは各データセットにわずかしかな存在しないため、トレーニングセットには品質の悪いサンプルより品質のよいサンプルの方が多く含まれている。これは、希少なクラスのサンプルは数多く入手できないという、分類における一般的な問題である。トレーニングセットに品質の悪いサンプルが少ないことは、ニューラルネットワーク実装[17]の「クラスの重み」パラメータを設定することによって補正した。このニューラルネットワーク実装は、クラス事前確率をトレーニングパラメータとして受け入れる。ここで使用しているスケールリングされた共役勾配法では、トレーニングに使われるネットワークエラーもエラー信号もクラスの重みを使って計算される。これにより、最適化を実行した

ときには大域的問題に対する最適解が生成されるが、より一般的でないクラスについては無理のないサンプリングが許される。

ニューラルネットワーク出力のクラスは、各トレーニングサンプルの(プローブ、ギャラリー)ペアについて計算される一致得点と不一致得点に基づいて決定される。一致得点が低くなる(および不一致の平均点が高くなる)のは、ペアの品質が低いためであることがわかっている。しかし、トレーニングの前に(プローブ、ギャラリー)のどちらの指紋画像の品質が低いかわからないため、プローブとギャラリーのどちらの指紋画像も同じクラスに割り当てられる。プローブとギャラリーの品質が同じであれば、サンプルが正しいクラスに割り当てられ、ニューラルネットワークは正しくトレーニングされる。しかし、プローブとギャラリーの品質が同じでなければ、(プローブ、ギャラリー)ペアのうち、品質のよい方の画像が間違ったクラスに割り当てられる。たとえば、品質のよい特徴点の数がひじょうに少なく、品質ゾーン4が小さい(つまり品質が低い)プローブ画像と、品質のよい特徴点の数が多く、品質ゾーン4が大きい(つまり品質がよい)ギャラリー画像を照合すると、正規化一致得点(方程式4)は低くなる。プローブとギャラリーがどちらも同じ(品質の低い)クラスに割り当てられることにより、ギャラリー画像(ペアのうち品質のよい方の画像)が明らかに間違ったクラスに割り当てられるためである。この問題は、ニューラルネットワーク実装の「パターンの重み」パラメータを調整することによって解決した。パターンの重みは、ニューラルネットワークがトレーニング中に最小化しようとする誤差値を計算するために使われる。トレーニングパターン(特徴)がネットワークを通過するとき生成された各パターンの拠出誤差は、そのパターンに対応するパターンの重みを掛けた上で累積誤差に追加される。

ここでは、ニューラルネットワークが適切にトレーニングされるように、各対象の(プローブ、ギャラリー)ペアのうち、品質のよい方の画像に低い重みを割り当てる必要がある。前述のように、トレーニングの前に(プローブ、ギャラリー)の相対的な品質はわからない。しかし、トレーニングのあとで各画像の予測されたクラスを使ってパターンの重みを調整し、調整されたパターンの重みを使ってトレーニングを繰り返すことができる。そこで、次のようにトレーニングを2回に分けて行った。最初に、すべての特徴ベクトルに同じ重みを与えて(すべての重みを1にして)ニューラルネットワークをトレーニングした。トレーニングが完了すると、ニューラルネットワークはトレーニングセットに含まれる各画像の品質のクラスを予測する。各対象の2つの画像(プローブ、ギャラリー)に関するこの予測品質に基づいて各画像に新しいパターンの重みを割り当て、新しい調整されたパターンの重みを使ってトレーニングを繰り返す。パターンの重みは、以下に示す方法で調整する。

トレーニングセット内の各対象(人)について、

$q_{\text{probe}} = q_{\text{gallery}}$  であれば、プローブのパターンの重みを0.5に設定し、ギャラリーのパターンの重みを0.5に設定する。

$q_{\text{probe}} < q_{\text{gallery}}$  であれば、プローブのパターンの重みを0に設定し、ギャラリーのパターンの重みを1に設定する。

$q_{\text{probe}} > q_{\text{gallery}}$  であれば、プローブのパターンの重みを1に設定し、ギャラリーのパターンの重みを0に設定する。

品質は5が最も低く、1が最も高いことに注意すること。

ニューラルネットワークに入力されるすべての特徴ベクトルが同じ範囲内にあることを保証するため、データセットDOS-C、DHS2-C、BEN、DHS10、およびTXDPSのすべての特徴ベクトルの大域的な平均と大域的な標準偏差によって、特徴ベクトルを正規化する。

ここでは、NIST[6]で行われたSDKテストのベンダーA～L、N、およびVTBの類似度得点を使って、ここで提案している指紋画像の品質測定基準をテストした。データセットVISIT\_POEに対するベンダーFの結果はこのレポートの本文で報告しており、ほかのすべてのデータセットに対するほかのすべてのベンダーの結果(合計280個のROC)は付録Aに掲載した。ニューラルネットワークのトレーニングにはVISIT\_POEデータセットのサブセットを使用しなかったことに注意されたい。

## 5 指紋画像品質の評価

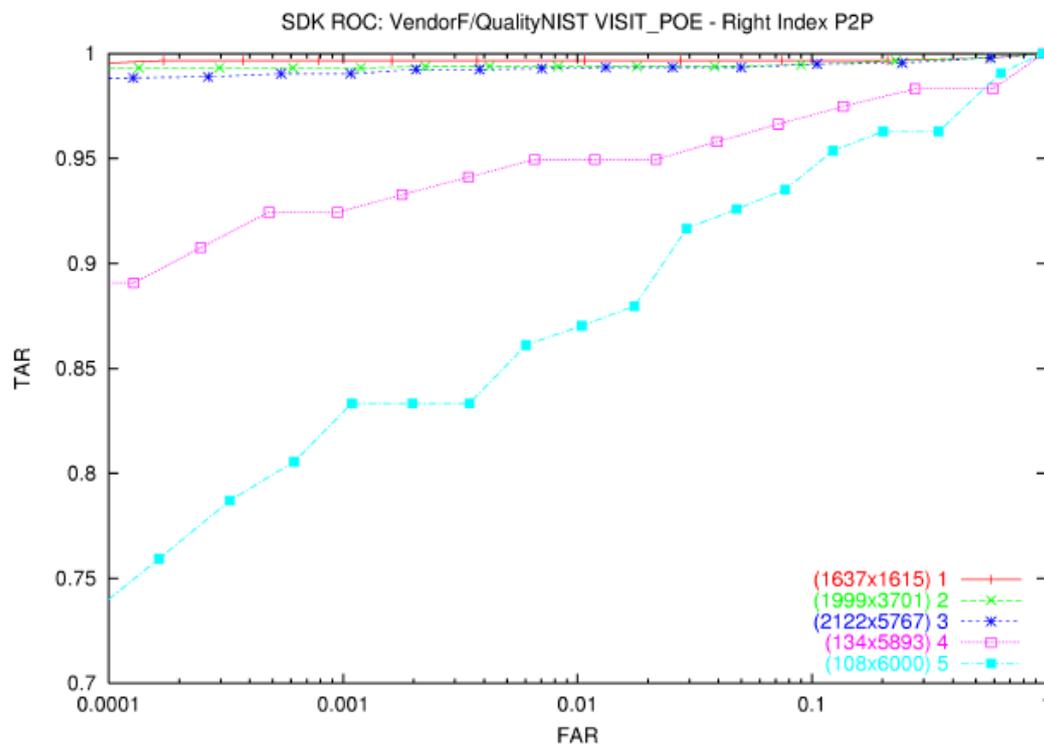
品質分析アルゴリズムの性能をテストするための指紋画像品質のベンチマークを得るのは困難である。画像品質測定基準をテストするのに最もよく使われる方法は、画像の視覚的な(したがって主観的な)評価に基づく方法である。しかし、データベースの規模が大きくなると、手作業での判定には莫大な時間とコストがかかる。現在のNIST指紋テストアーカイブの全体は、1千6百万人から収集したおよそ6千8百万個の指紋で構成されている。したがって、感じ取られた画像品質と照合性能を許容できるレベルで対応付ける客観的な評価基準が不可欠である。

ここで提案する評価基準は、画像品質の関数としての順位ROCである。テストデータセット内の各対象者のプローブとギャラリーの指紋ペアに対する品質値 $q$ を測定する。品質は、2つの値の小さい方を使って定義される。品質値が $K$ 個の整数値を取るとすると、 $K$ 個のROC特性が生成される。 $k$ 番目の品質値を $q_k$ とすると、 $K$ の不一致分布 $N_k(s)$ および $K$ の一致分布 $M_k(s)$ の計算に使われる類似度得点は、方程式6および8によって画像ペアの品質が $q_k$ になる場合の得点である。値 $K$ の選択は、設計上の問題である。 $K$ は少なくとも2である。その場合は、画像を品質がよいとして受け入れるか品質が悪いとして拒否するかの二者択一になる。ここでは、 $K$ を5とした。画像品質を5レベルにするのは、1つの妥協案である。画像品質にひじょうに敏感な照合プログラム(たとえば、SDK照合プログラムA[6])を完全に特徴付けるには、8～10レベルが必要である。画像品質にあまり敏感でない照合プログラム(たとえば、SDK照合プログラムHおよびI)であれば、画像品質を3レベルにしても十分に特徴付けることができるだろう。ここでは、品質が中～高レベルの照合プログラム(たとえば、SDK照合プログラムF)にとって適切な5レベルを選択した。

図11-aおよび11-bに、データセットVISIT\_POEの右手および左手人差し指の品質分析とSDKベンダーFの類似度得点の関数としての順位ROC性能を示す。これらの図は、ここで提案している画像品質が性能の順位統計として機能していることを示している。ただし、ニューラルネットワークのトレーニングにはVISIT\_POEデータセットのサブセットを使用しなかったことを述べておきたい。

結果を示す図(図11-24)は、それぞれ3つのセクションで構成される。図の1つ目のセクションには、画像品質レベル1～5の順位付けしたROC曲線が含まれる。異なる品質レベルに対する性能の違いを調べるには、5つ(品質のクラスごとに1つ)のROCの結果を結合する必要がある。結果を結合する1つの方法は、各FARに対するTAR(True Accept Rate, False Accept Rate)の偏差を測定することである。このためには、各プローブセットの動作しきい値(つまり、異なる品質レベル)を調整する必要がある。多くのアプリケーションでは、これは実現不可能か、または望ましくない。しかし、異なるシステム間に同じしきい値を設定することは不可能なので、異なるシステムの性能を比較するためにはこれが適切な手法である。システムの性能に対する品質の影響を調べるには、すべてのギャラリーとプローブセットに対して1つの動作しきい値を設定しなければならない。さまざまなレベルの品質に対してしきい値を固定することは、1つのアプリケーションに関してしきい値を一度だけ設定する運用状況に対応している。この動作しきい値は、ある集団に対する特定の他人受入率(通常は1%)に基づいて選択される。各図の2つ目のセクションには、この集団の動作しきい値とそれに対応するFAR値とTAR値、および動作しきい値で計算された各画像品質レベルのFAR値とTAR値を示す表が含まれる。各図の3つ目のセクションには、2つ目のセクションに示された値のグラフが含まれる。

具体的には、品質が異なるプローブセットの固定された動作しきい値での他人受入率と本人受入率が示される。赤い四角は、固定された動作しきい値でのデータセット全体の他人受入率と本人受入率を示している。



Vendor F – VISIT_POE – Right index threshold=350 (far,tar)=(0.0119418,0.991333)					
quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0037	0.0083	0.0131	0.0216	0.0477
TAR	0.997	0.994	0.993	0.9496	0.926

(FAR,TAR)at fixed threshold - Vendor F - VISIT\_POE - Right Index

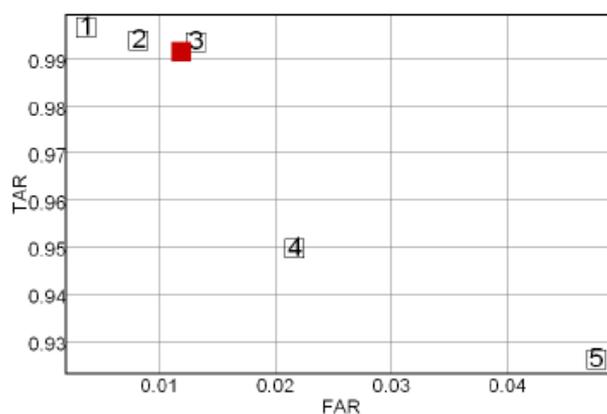


図11-a. ROCに対する品質の影響－ベンダーF－データセットVISIT\_POE－右手人差し指

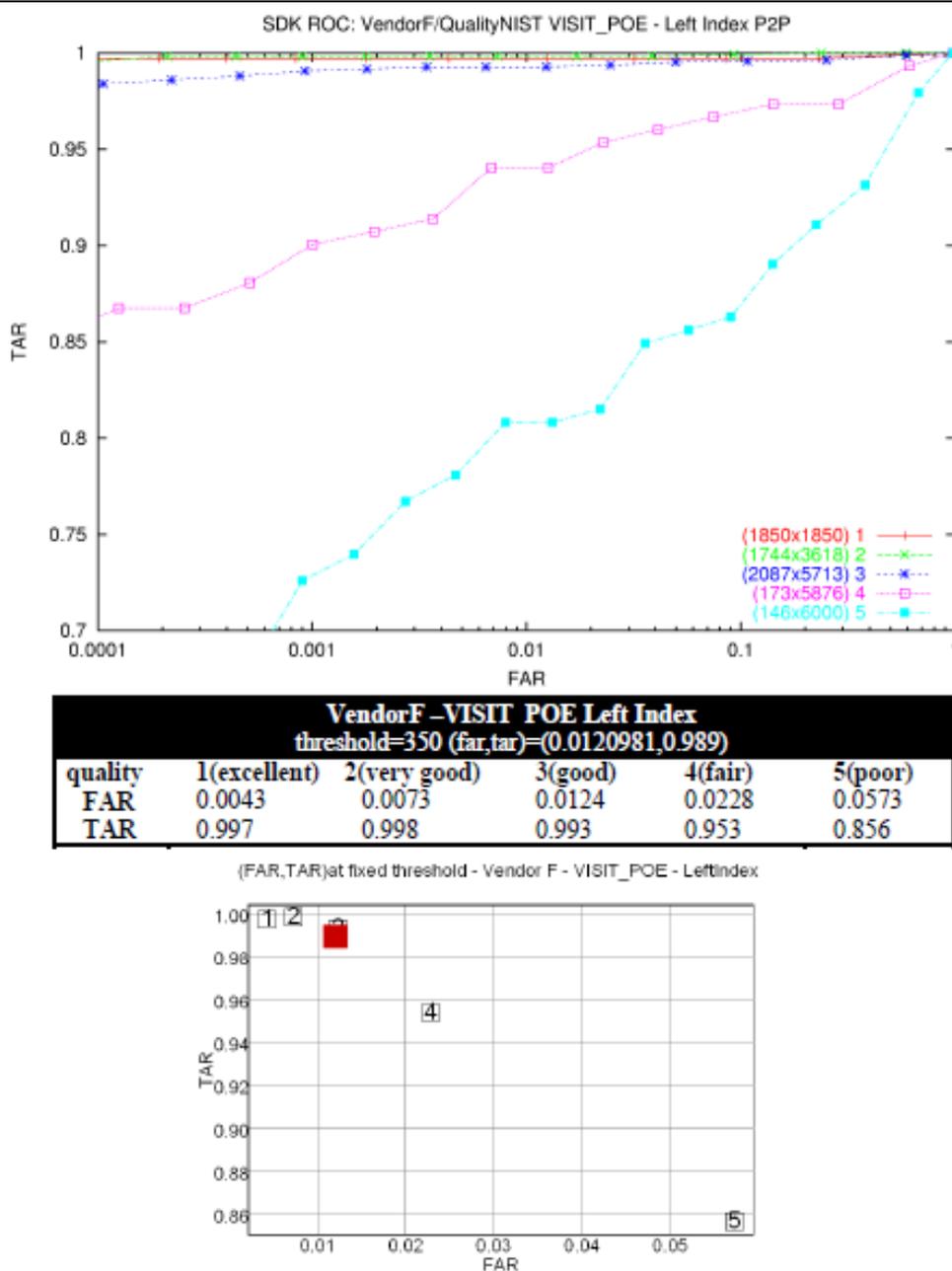


図11-b. ROCに対する品質の影響－ベンダーF－データセットVISIT\_POE－左手人差し指

品質測定基準の望ましい特徴は、ほかの照合プログラムやほかのデータセットに、いかに問題なく一般化できるかである。このシステムがほかの照合プログラムにいかに問題なく一般化できるかを調べるため、SDKベンダーA～L、N、およびVTBの類似度得点を使って、1.2で説明したすべてのデータセットに対してここで提案している指紋画像測定基準をテストした。データセットVISIT\_POEの結果は図11～24に示し、すべてのベンダー（14ベンダー）とすべてのデータセット（20データセット）の結果（合計208個のROC）は付録Aに掲載した。これらのROCは、この指紋品質測定基準が何の問題もなくほかの照合プログラムやほかのデータセットに一般化できることを示している。

各ROCのあとに、同じ動作基点（つまり固定されたしきい値）による各品質の（他人受入率、本人受入率）ペアの表とグラフを示す。対応するデータセットのしきい値と（他人受入率、本人受入率）も報告する。赤い四角は、各図の中央の表に指定された動作しきい値でのデータセット全体の他人受入率と本人受入率を示している。品質が低いほど、性能もさがる（つまり、本人受入率がさがり、他人受入率が上がる）ことが予想される。図11～24は、ベンダーF、C、G、I、およびJについて、品質1および2（「秀」と「優」）の指紋の他人受入率と本人受入率がほぼ必ず赤い四角の左側にあることを示している。これは、品質1（秀）および品質2（優）のプローブセットの照合性能がデータセット全体より高いことを意味する。品質3、4、5（良、可、不可）の他人受入率と本人受入率は、通常、赤い四角の右側にあり、データセット全体より本人受入率が低く、他人受入率が高いことを示している。

品質が低いほど、すべての照合プログラムで本人受入率はさがる。品質が高いほどFARがあがることも確認されたい。ベンダーF、C、G、J、およびIはにこのような性質（品質が低いほどTARがさがり、FARがあがる）がみられる。ベンダーA、B、N、およびVTBでは、品質が低いほど他人受入率はさがる。ベンダーD、E、H、K、およびLでは、品質の違いによるFARのパターンはない。品質が低いほど他人受入率が上がることは、品質の低い指紋画像の脆弱性を示している。

ベンダー照合プログラムの性能の、この文書で定義した画像品質に対する感度には興味深い傾向がいくつかある。品質レベル1のデータですべてのベンダーを調べると、TARが99.6%を超えるベンダーはごく一部である。これは、この文書で定義した画像品質とは別に、およそ0.4%の誤差を生む誤差のソースがあることを示している。

同時に、この動作しきい値での各ベンダーのFARは、ベンダーNの2.8%からベンダーIの0.02%までの範囲にある。これは、SDKテストでの精度の順位が比較的低いシステムでは、システムのFARとTARが関連しにくいことを示している。

品質1の指紋に関するベンダーのFARが0.1%未満であれば、ほとんどの場合、そのベンダーは品質1、2、および3の指紋に対して同じような反応を示す。このことから、精度がひじょうに高いシステムの性能を完全に特徴付けるのに必要な画像品質のレベルは5未満であることがわかる。

最後に、SDKレポートの結果を検討したところ、上位5つのシステム（SDKシステムH、I、J、F、およびG）では、VISIT\_POEデータの画像品質3のTARがSDKの順位に関する優れた総合的な予測材料であることがわかった。異なるデータセットに関する各種のSDKベンダーシステムの順位（[6]より抜粋）を表6に示す。固定された動作しきい値での品質3のTARに基づく上位5ベンダーの順位を表7に示す。DOS-CおよびDHS2-Cの人差し指のデータセットの順位と品質3のTARの順位では、HとFの位置が異なるだけである。この順位とほかのSDKデータセットとの違いは、BEN、DHS10、TXDPS、OHIO、およびSD29データセットでテストされた平面押捺指紋対回転押捺指紋の照合方法が原因のひとつである。

DOS-C/DHS2-C	BEN/DHS10/ TXDPS	すべて
システム I	システム H	システム H
システム H	システム I	システム I
システム F	システム J	システム J
システム J	システム F	システム F
システム G	システム G	システム G
システム D	システム D	システム D
システム C	システム K	システム C
システム K	システム C	システム K
システム L	システム L	システム L
システム B	システム VTB	システム A
システム VTB	システム A	システム VTB
システム A	システム B	システム B
システム E	システム E	システム E

表6. 各種データセットのFAR0.01%でのSDKの順位([6]より抜粋)

品質3の 左手 人差し指の 順位	VISIT_POEの 右手 人差し指の TAR	VISIT_POEの 左手 人差し指の TAR
システムI	0.993	0.996
システムF	0.993	0.993
システムH	0.993	0.995
システムJ	0.986	0.990
システムG	0.984	0.988

表7. US\_VISIT\_POEの左手人差し指の指紋を使った画像品質3のTARに関する各種ベンダーの画像品質順位。しきい値およびTARの値は、システムFについては図11に基づき、そのほかのシステムについては図17~20に基づく。

VISIT\_POEデータセットの品質の分布を表8に示す。データセットDOS-C、DHS2-C、DHS10、BEN、TXDPS、およびVISIT\_POE\_BVAの分布は、付録Aに掲載した。

	品質	テストセ ットのサ ンプル数	全体に 占める 割合	品質分布の比較				
				q1	q2	q3	q4	q5
右手人差し指	1(秀)	1637	27.28	28	32	35	3	3
	2(優)	1999	33.32	34	30	36	4	4
	3(良)	2122	35.37	27	29	35	3	3
	4(可)	134	2.23	2	2	2	2	2
	5(不可)	108	1.8	1	1	1	1	1
左手人差し指	1(秀)	1850	30.83	28	32	35	3	3
	2(優)	1744	29.07	34	30	36	4	4
	3(良)	2087	34.78	27	29	35	3	3
	4(可)	173	2.88	2	2	2	2	2
	5(不可)	146	2.43	1	1	1	1	1

表8. データセットVISIT\_POEの品質の割合

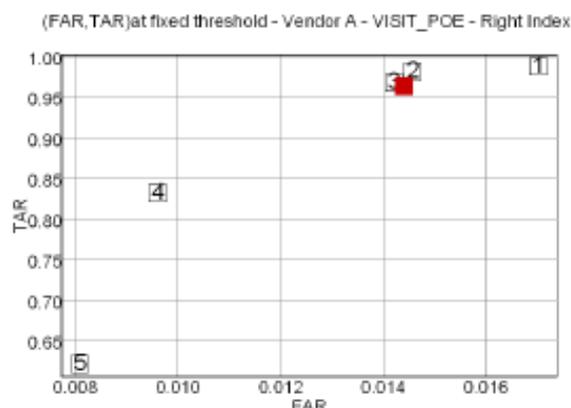
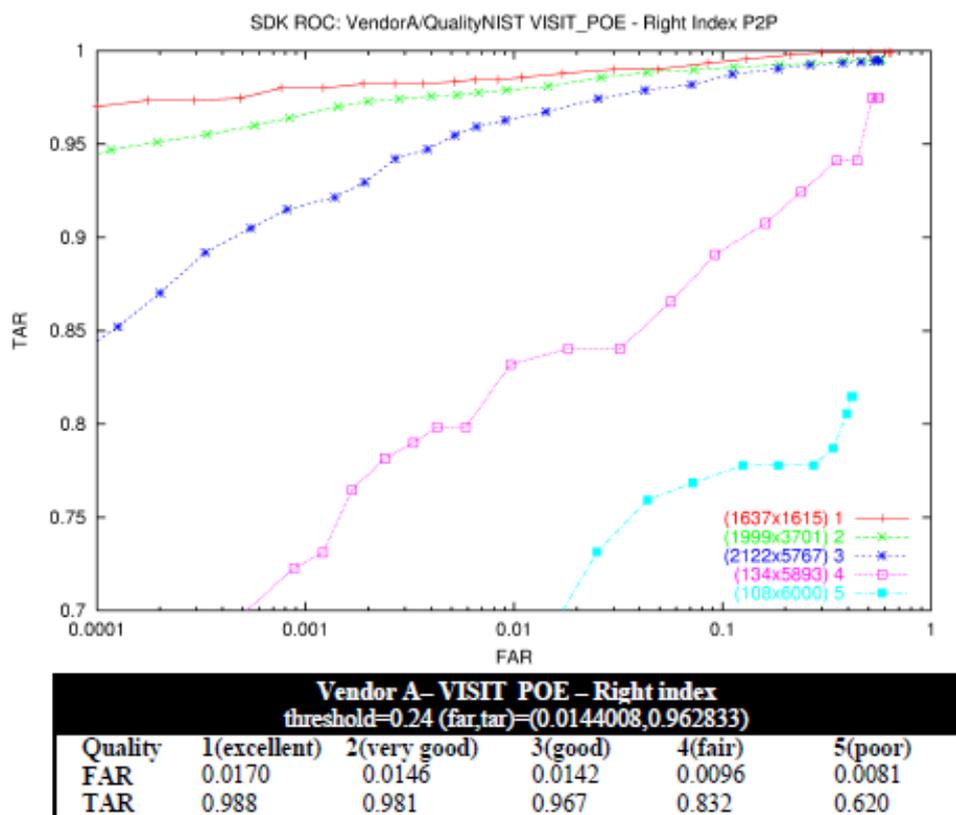


図12-a. ROCに対する品質の影響－ベンダーA－データセットVISIT\_POE－右手人差し指

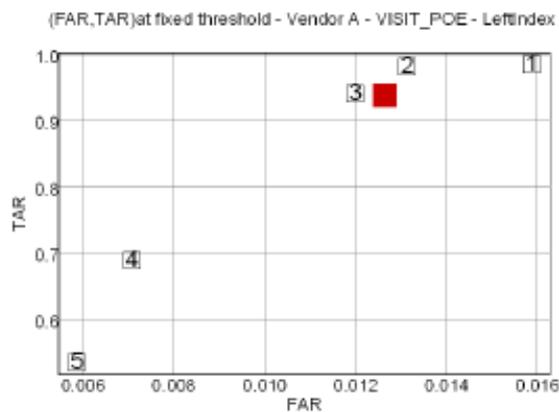
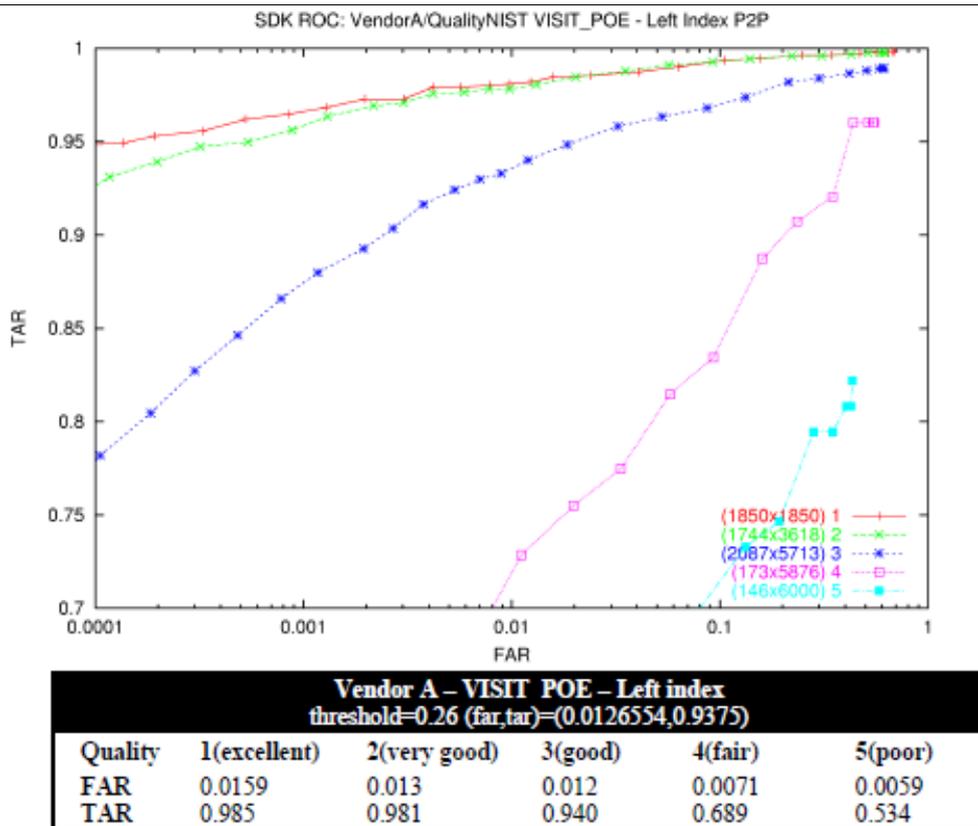


図12-b. ROCに対する品質の影響－ベンダーA－データセットVISIT\_POE－左手人差し指

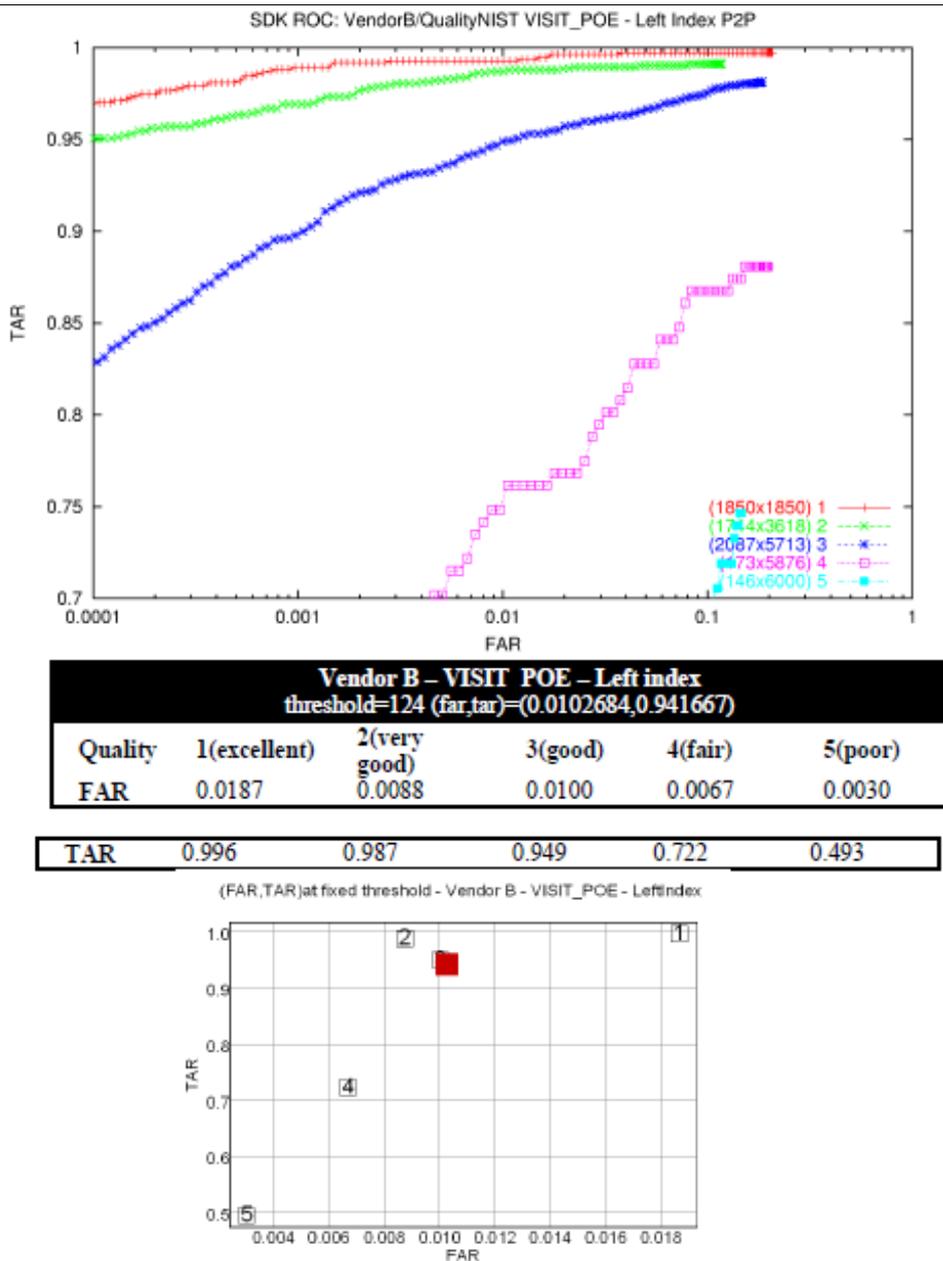
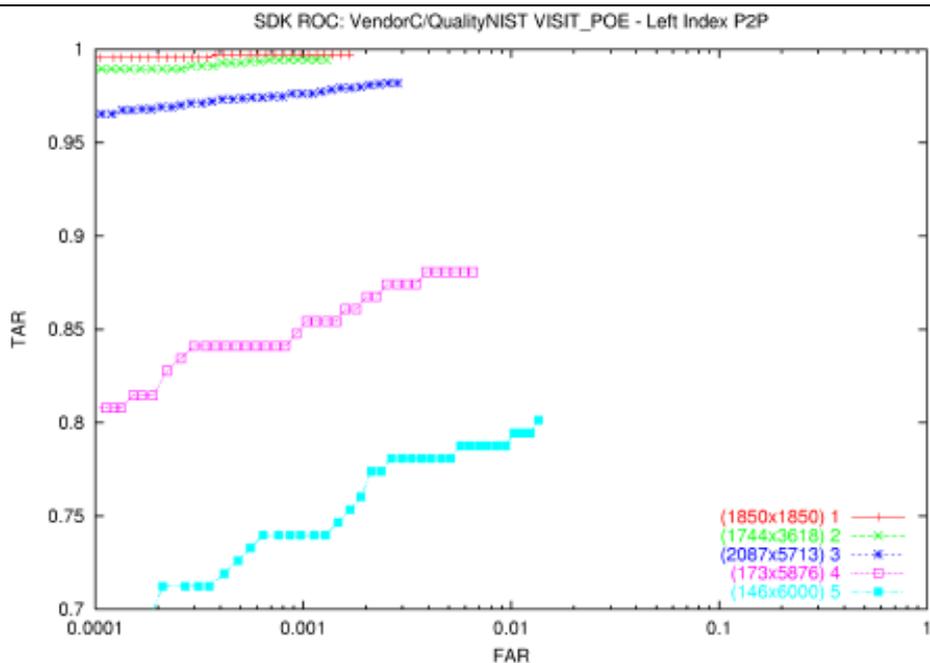


図13-a. ROCに対する品質の影響－ベンダーB－データセットVISIT\_POE－左手人差し指



Vendor C - VISIT POE - Left index threshold=110 (far,tar)=(0.001,0.975)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.00066	0.00045	0.00099	0.00225	0.00511

TAR	0.997	0.993	0.976	0.867	0.781
-----	-------	-------	-------	-------	-------

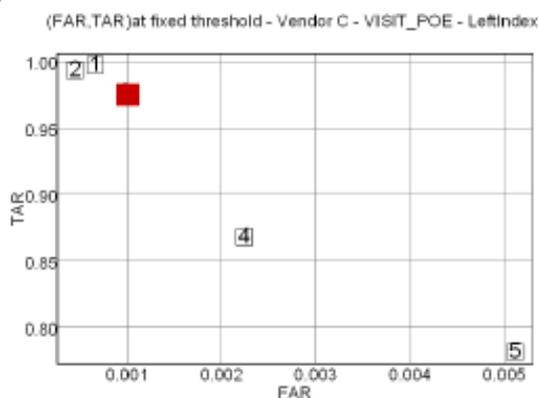
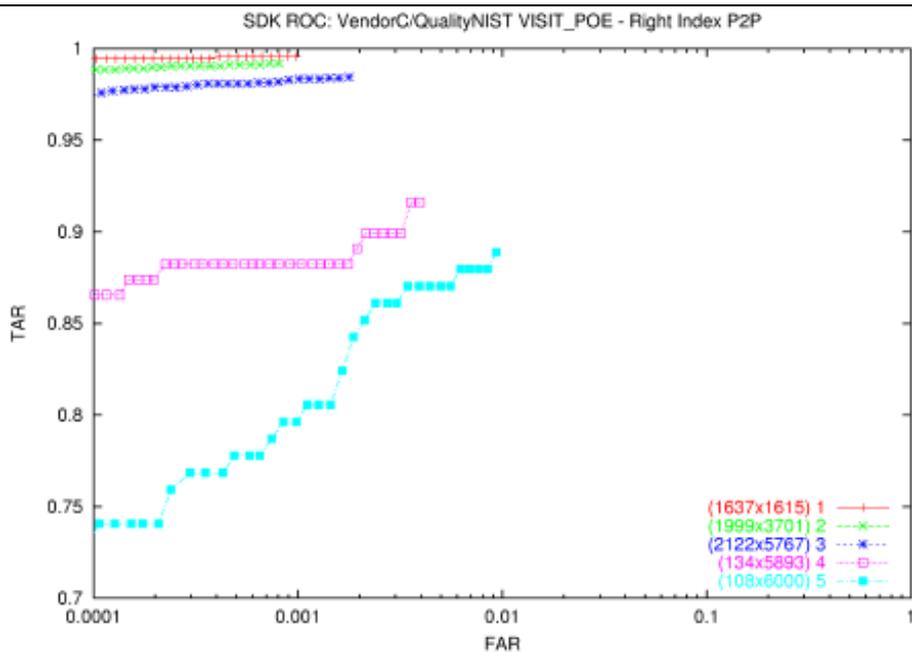


図14-a. ROCに対する品質の影響－ベンダーC－データセットVISIT\_POE－左手人差し指



Vendor C - VISIT_POE - Right index threshold=50 (far,tar)=(0.001,0.982)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.00067	0.00052	0.00114	0.00261	0.00624

TAR	0.995	0.991	0.983	0.899	0.880
-----	-------	-------	-------	-------	-------

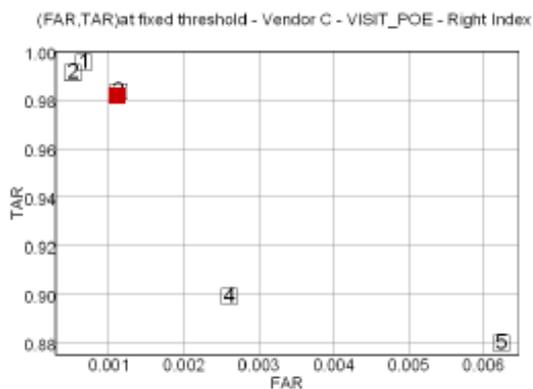
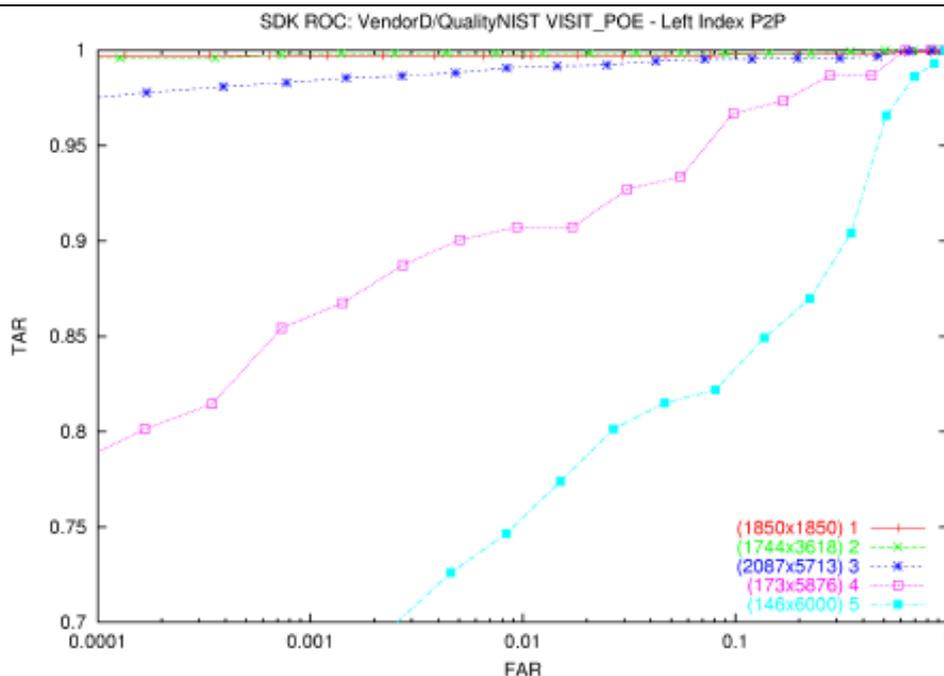


図14-b. ROCに対する品質の影響—ベンダーC—データセットVISIT\_POE—右手人差し指



Vendor D - VISIT_POE - Left index					
threshold=600 (far,tar)=(0.016,0.984)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0179	0.0206	0.0145	0.0094	0.0150
TAR	0.997	0.998	0.992	0.907	0.774

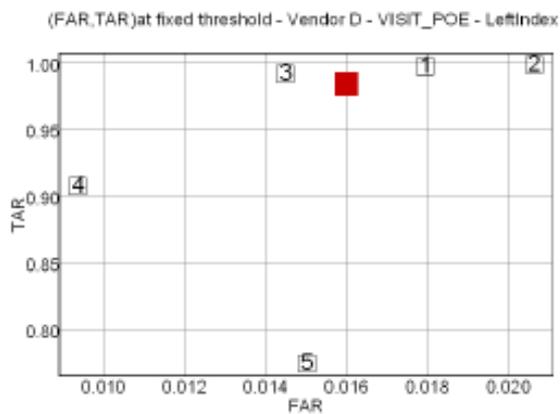
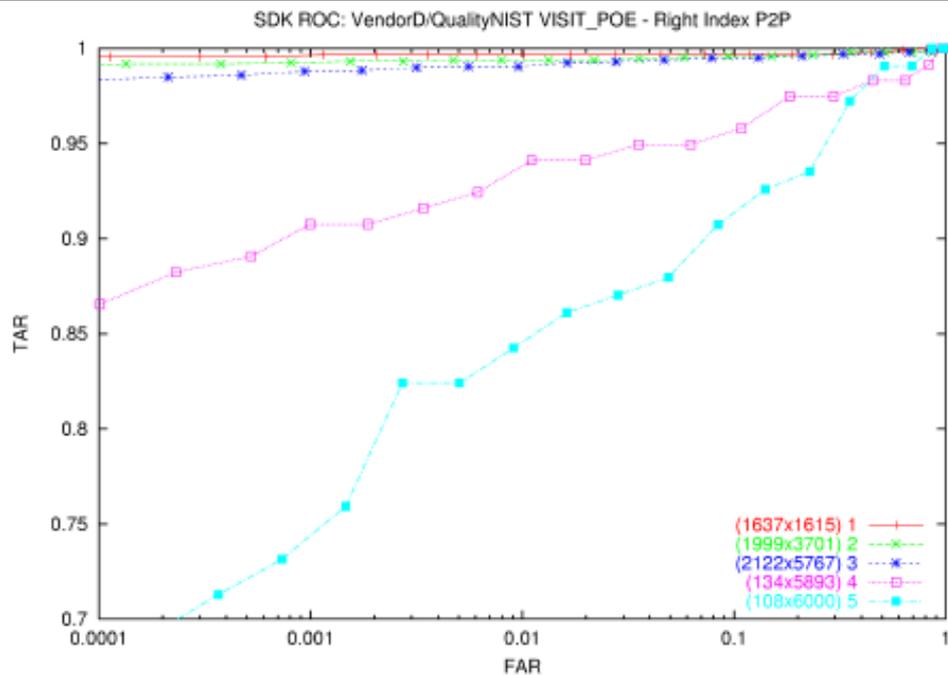


図15-a. ROCに対する品質の影響－ベンダーD－データセットVISIT\_POE－左手人差し指



Vendor D - VISIT_POE - Right index					
threshold=650 (far,tar)=(0.01,0.986)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.010	0.013	0.0096	0.006	0.0091
TAR	0.997	0.994	0.990	0.924	0.8426

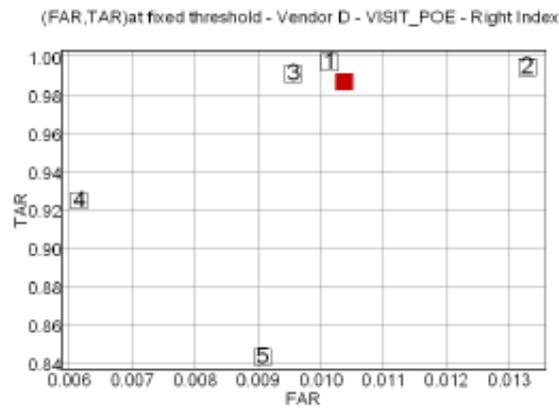


図15-b. ROCに対する品質の影響－ベンダーD－データセットVISIT\_POE－右手人差し指

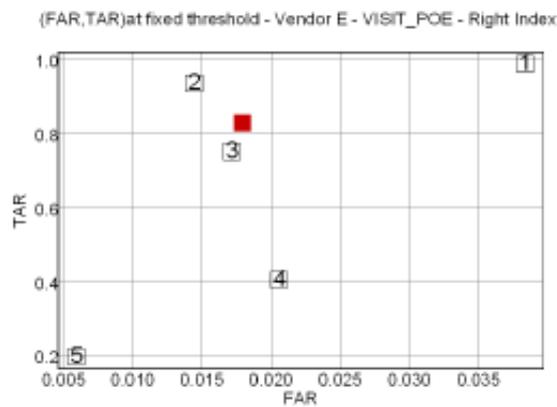
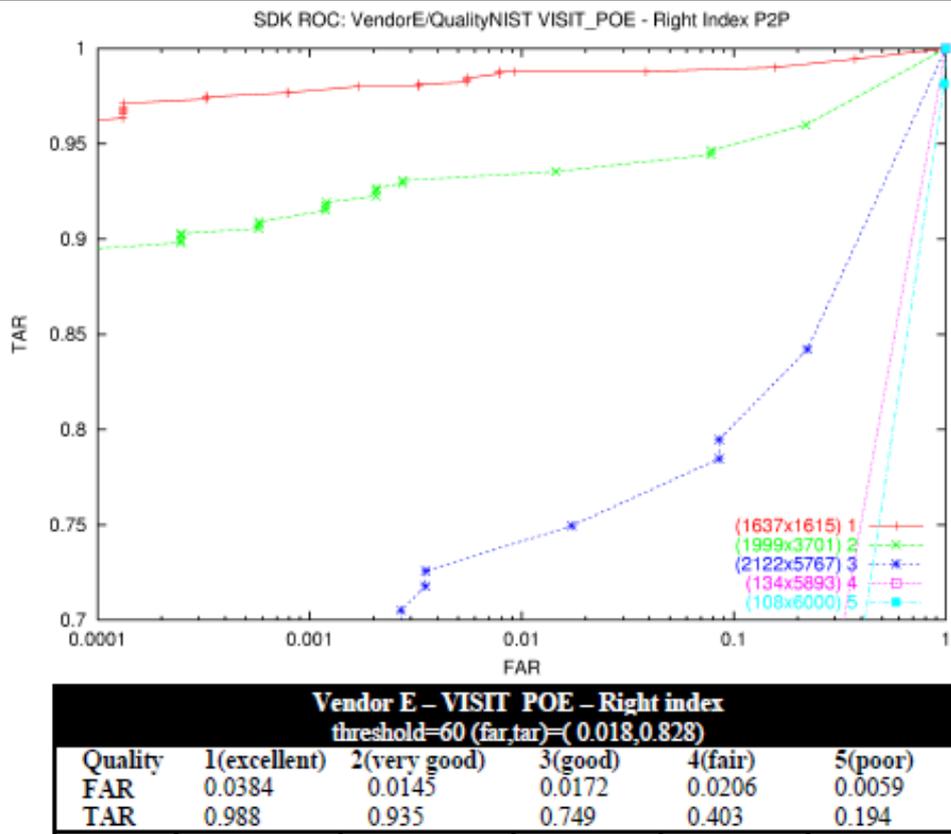


図16-a. ROCに対する品質の影響－ベンダーE－データセットVISIT\_POE－右手人差し指

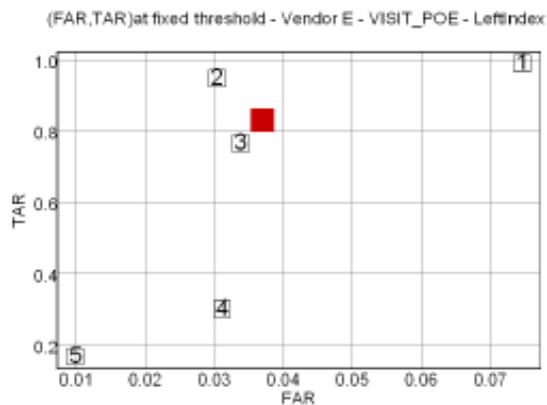
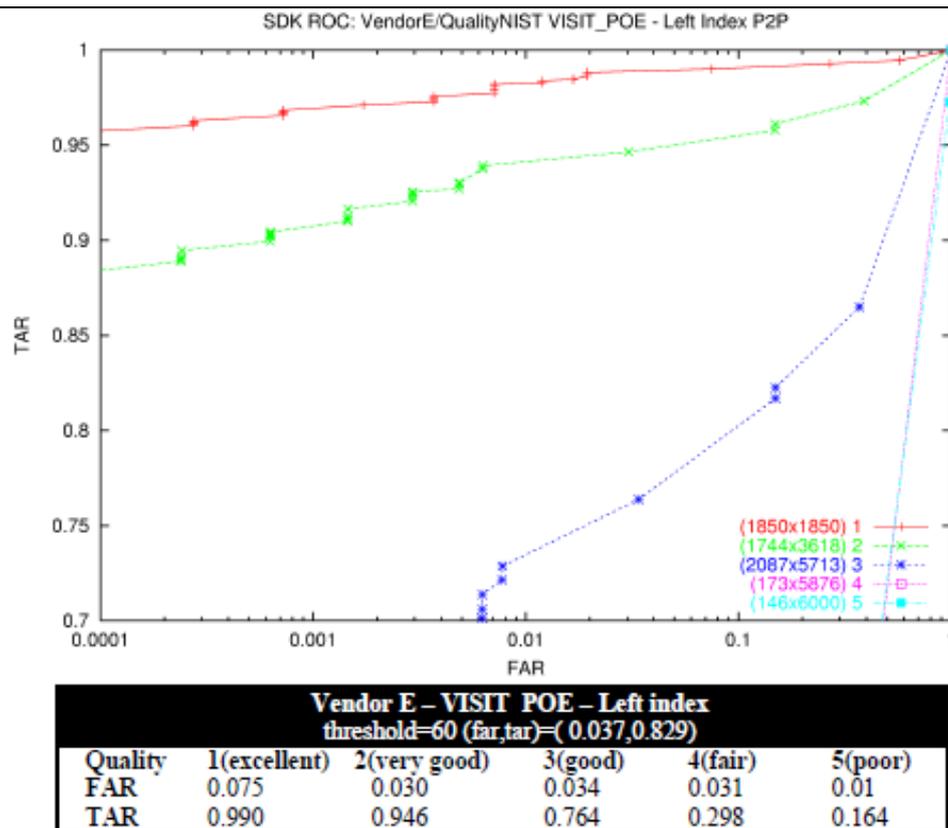


図16-b. ROCに対する品質の影響－ベンダーE－データセットVISIT\_POE－左手人差し指

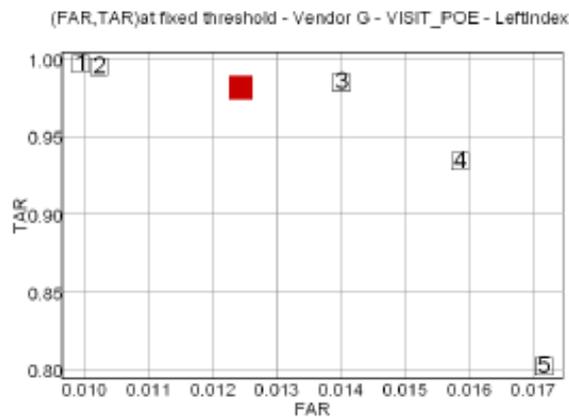
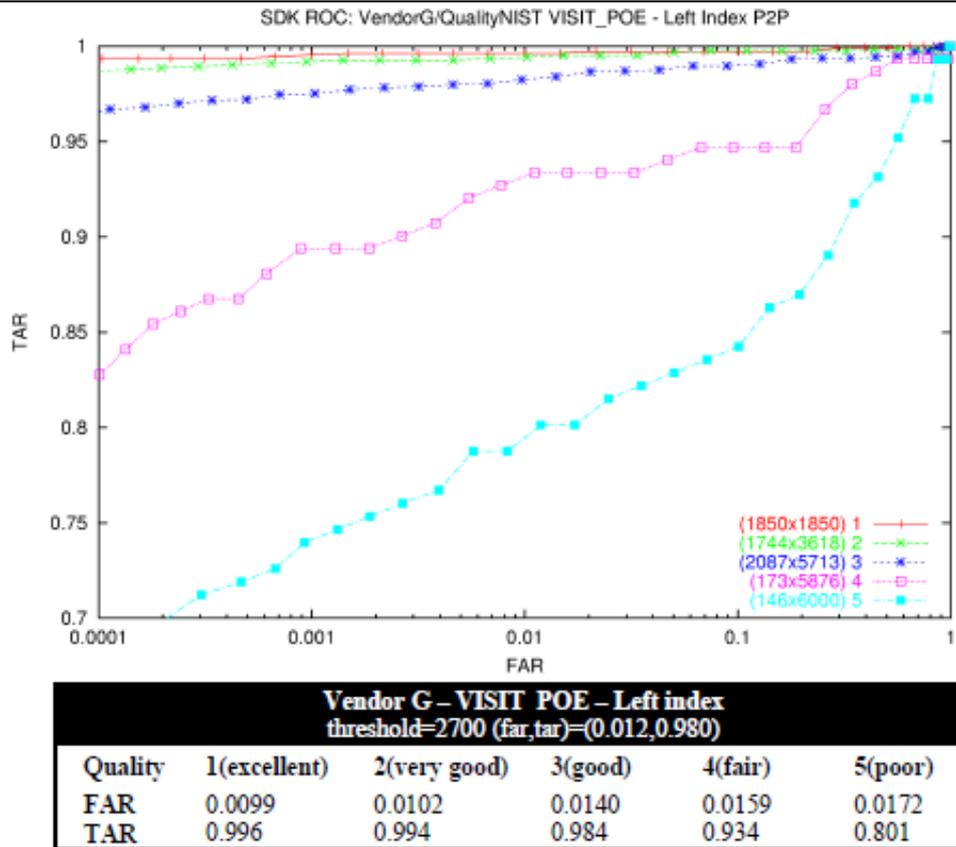


図17-a. ROCに対する品質の影響－ベンダーG－データセットVISIT\_POE－左手人差し指

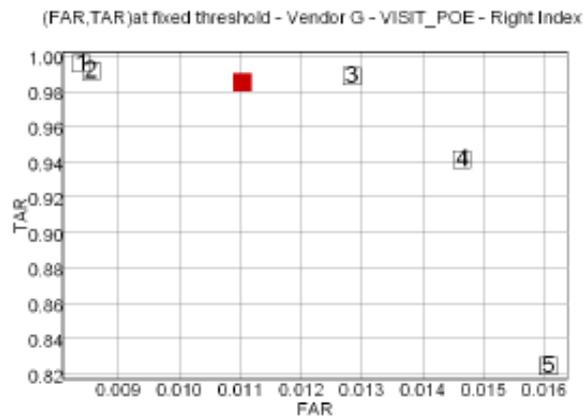
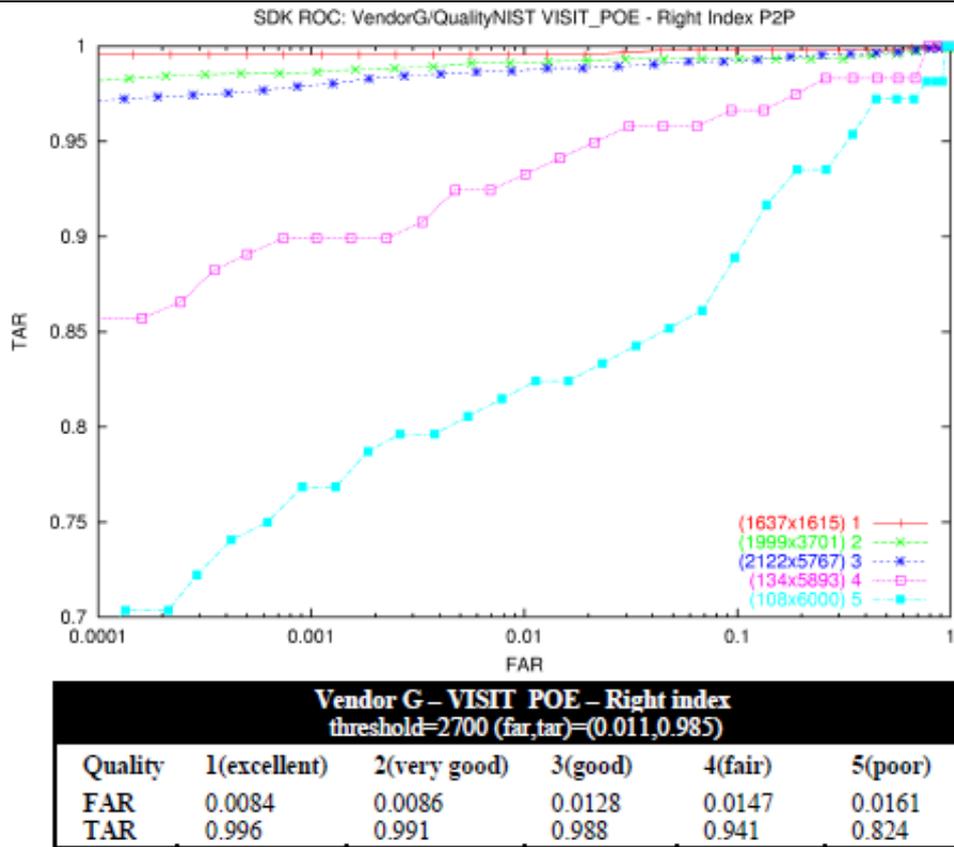
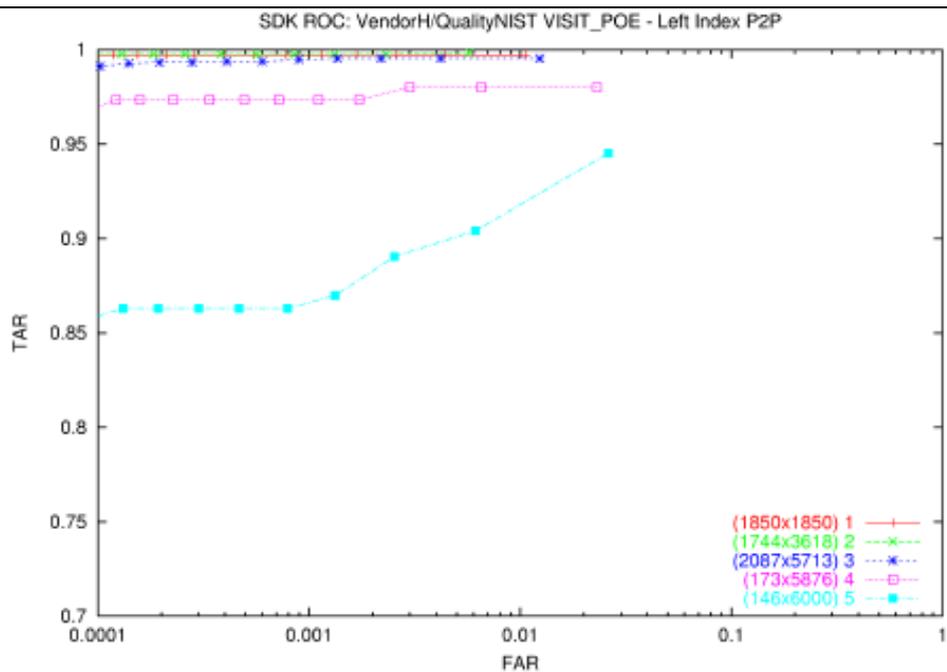


図17-b. ROCに対する品質の影響－ベンダーG－データセットVISIT\_POE－右手人差し指



Vendor H - VISIT POE - Left index threshold=40 (far_tar)=(0.001,0.991)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0017	0.0008	0.0014	0.0017	0.0013
TAR	0.997	0.998	0.995	0.974	0.870

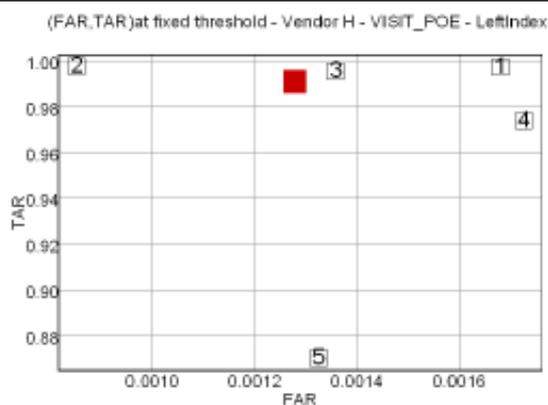
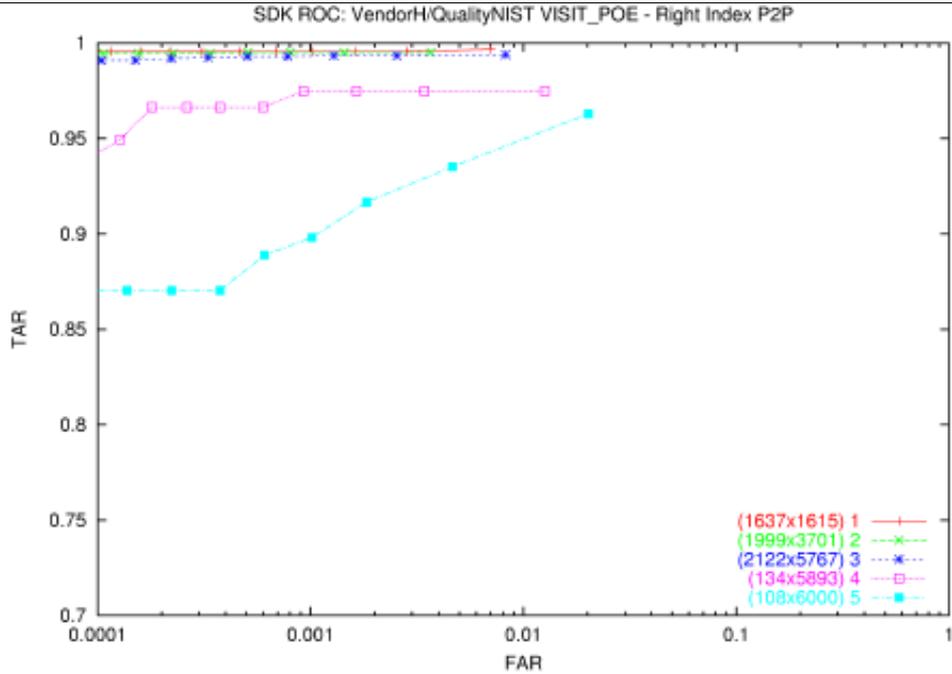


図18-a. ROCに対する品質の影響—ベンダーH—データセットVISIT\_POE—左手人差し指



Vendor H - VISIT POE - Right index					
threshold=30 (far,tar)=(0.001,0.992)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0016	0.00080	0.00129	0.00164	0.00183
TAR	0.996	0.995	0.993	0.975	0.917

(FAR,TAR)at fixed threshold - Vendor H - VISIT\_POE - Right Index

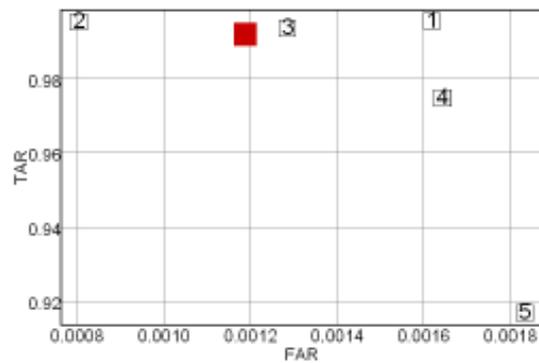
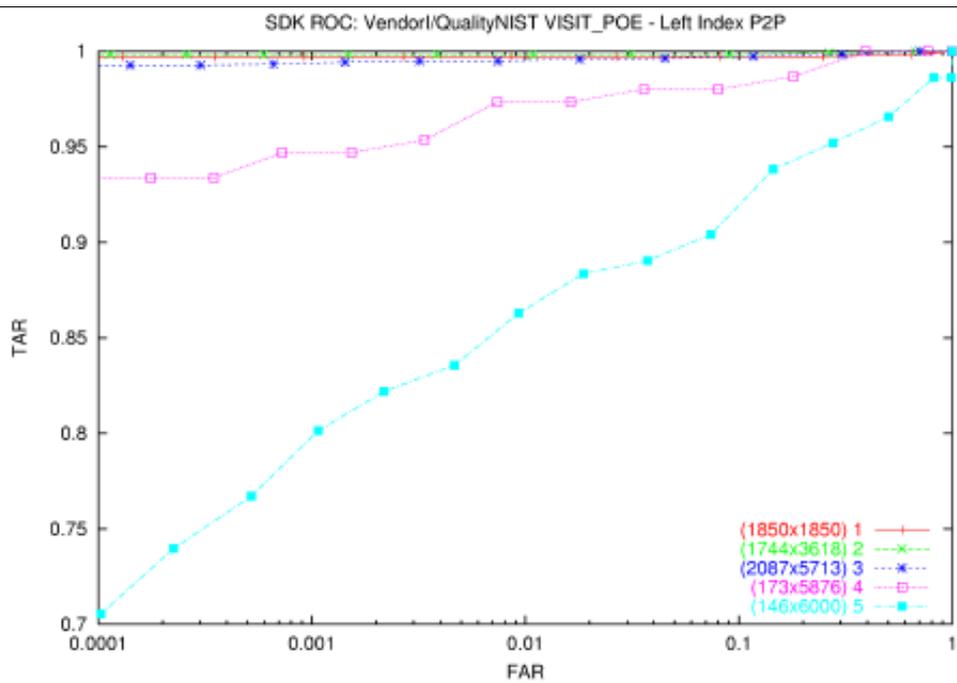


図18-b. ROCに対する品質の影響－ベンダーH－データセットVISIT\_POE－右手人差し指



Vendor I - VISIT POE - Left index threshold=400 (far,tar)=( 0.001,0.992)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.00058	0.00101	0.00191	0.00383	0.01129
TAR	0.998	0.997	0.996	0.972	0.880

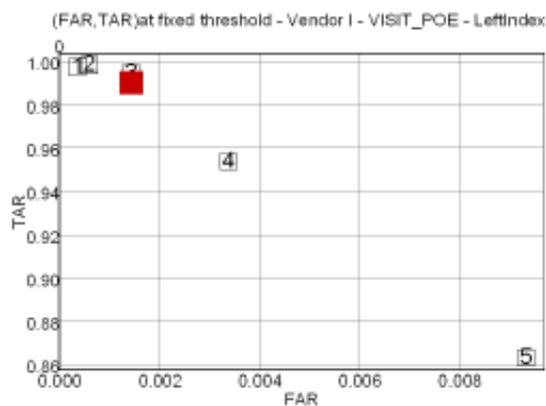
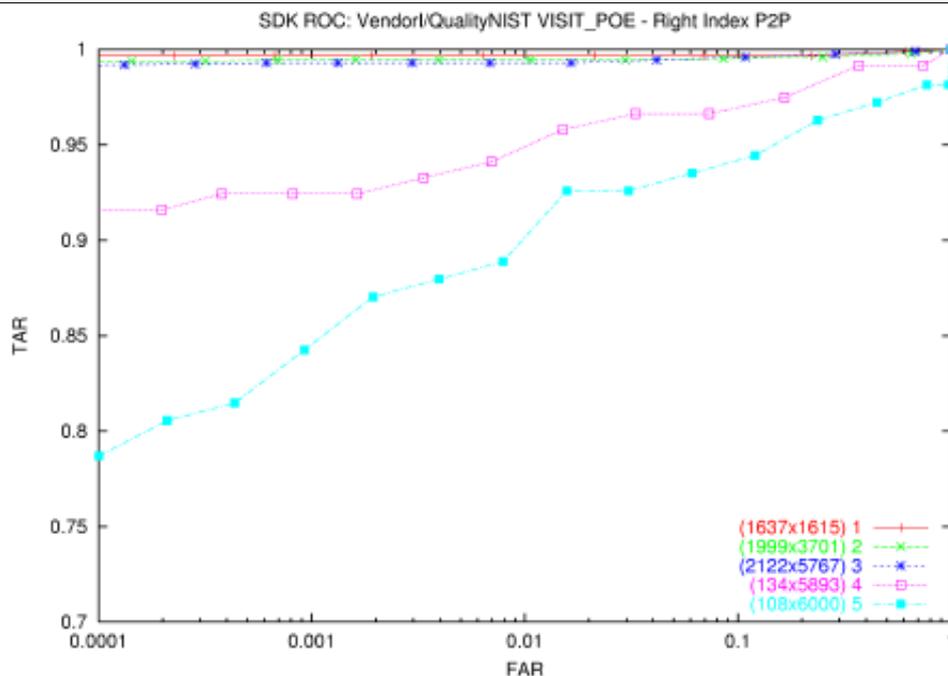


図19-a. ROCに対する品質の影響－ベンダーI－データセットVISIT\_POE－左手人差し指



Vendor I - VISIT POE - Right index threshold=450 (far,tar)=(0.001,0.99)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.00022	0.00069	0.00133	0.00336	0.00789
TAR	0.997	0.995	0.993	0.933	0.889

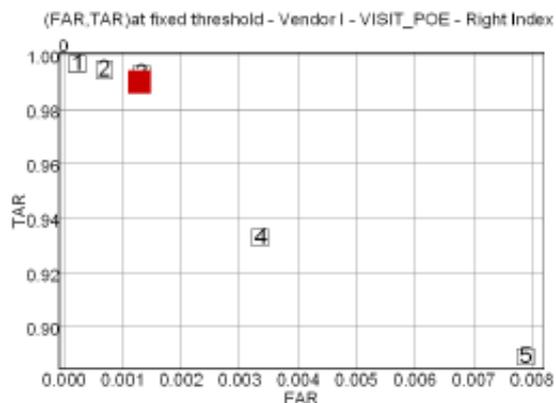
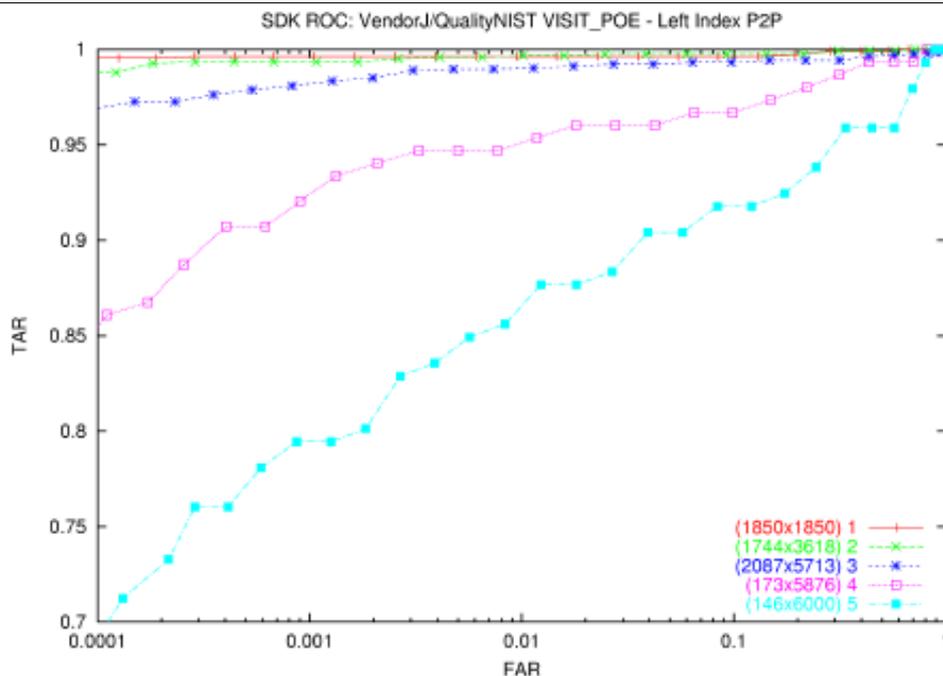


図19-b. ROCに対する品質の影響－ベンダーI－データセットVISIT\_POE－右手人差し指



**Vendor J - VISIT\_POE - Left index**  
threshold=2600 (far,tar)=(0.01,0.987)

Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0094	0.0101	0.0114	0.0118	0.0181
TAR	0.996	0.997	0.990	0.954	0.877

(FAR,TAR)at fixed threshold - Vendor J - VISIT\_POE - LeftIndex

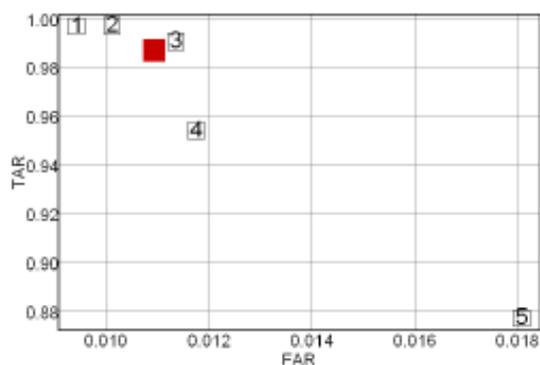
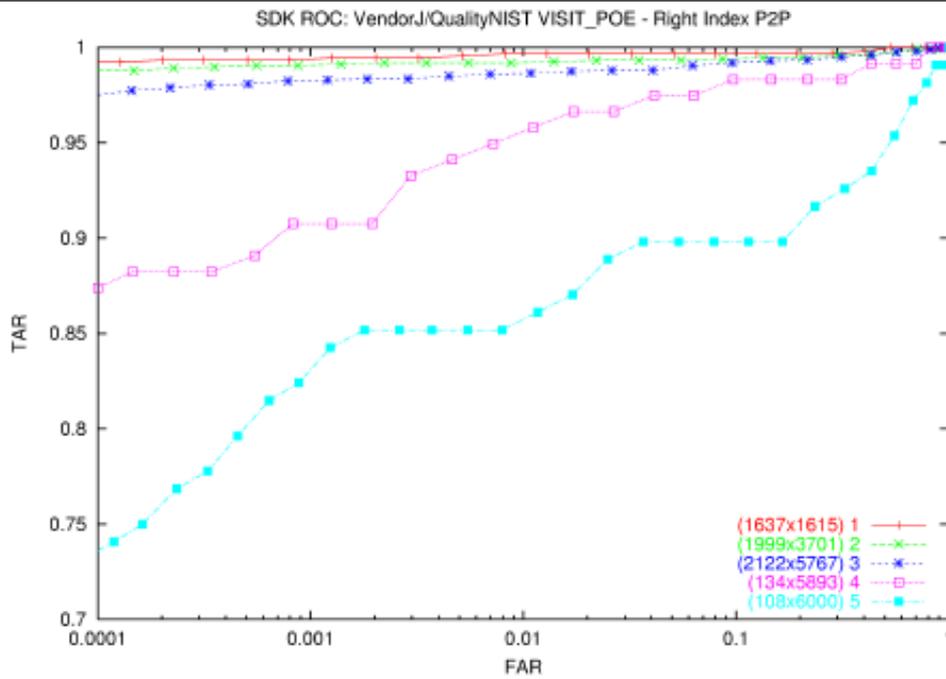


図20-a. ROCに対する品質の影響－ベンダーJ－データセットVISIT\_POE－左手人差し指



Vendor J - VISIT POE - Right index threshold=2600 (far,tar)=(0.01,0.987)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0082	0.0088	0.0108	0.0111	0.0170
TAR	0.997	0.992	0.986	0.958	0.870

(FAR,TAR)at fixed threshold - Vendor J - VISIT\_POE - Right Index

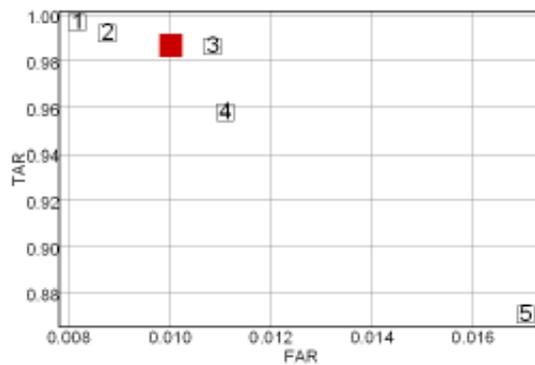
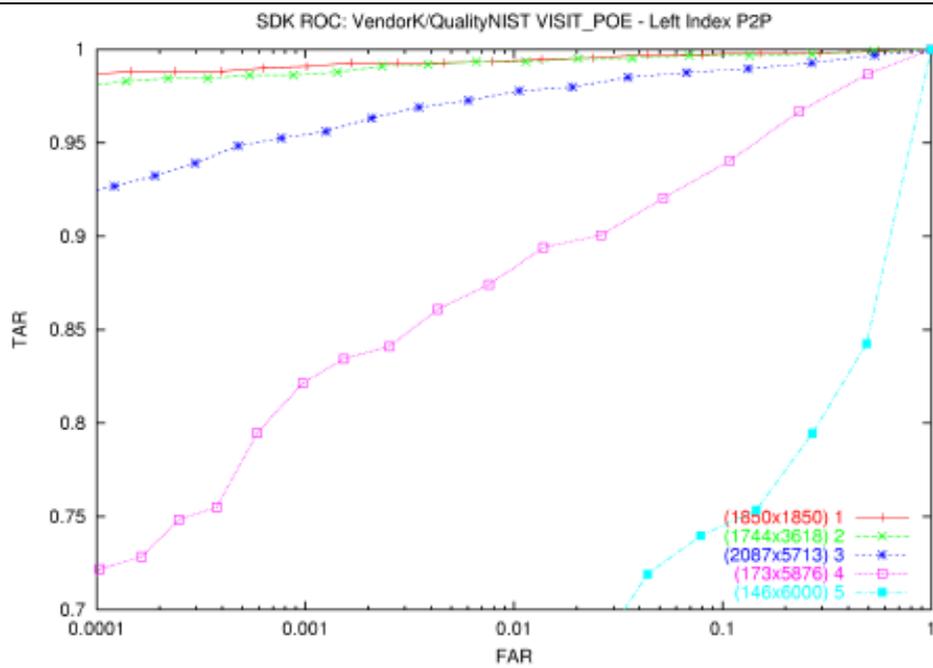


図20-b. ROCに対する品質の影響－ベンダーJ－データセットVISIT\_POE－右手人差し指



Vendor K - VISIT_POE - Left index					
threshold=14 (far,tar)=(0.011,0.969)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0136	0.0114	0.0106	0.0076	0.0147
TAR	0.995	0.994	0.978	0.874	0.651

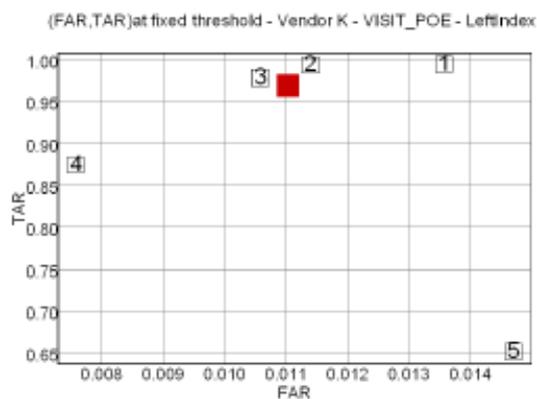
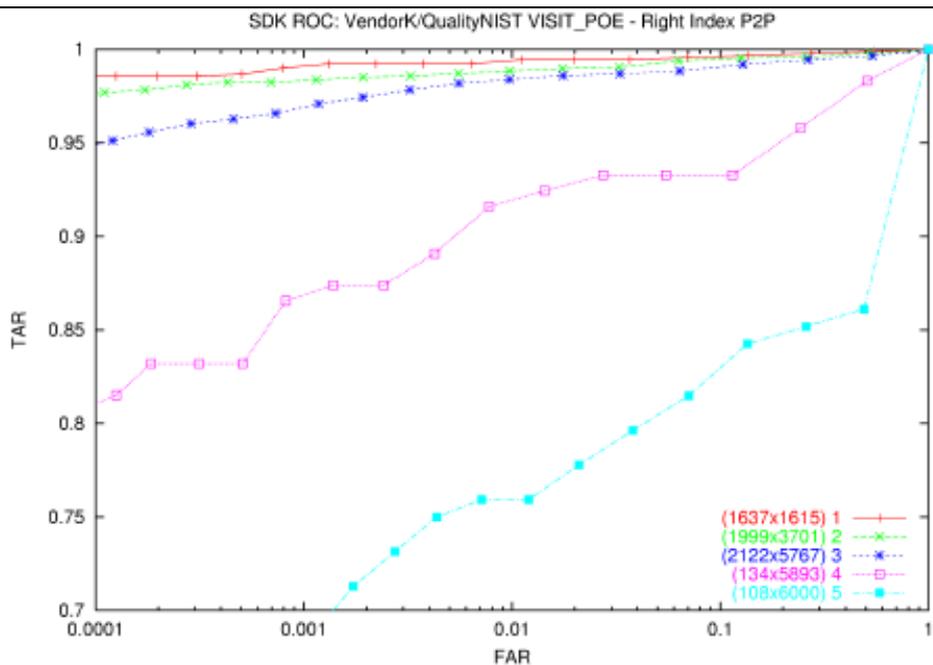


図21-a. ROCに対する品質の影響－ベンダーKデータセットVISIT\_POE－左手人差し指



Vendor K – VISIT POE – Right index threshold=12 (far,tar)=(0.018,0.980)					
quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0199	0.0175	0.0176	0.0143	0.0210
TAR	0.994	0.990	0.986	0.924	0.778

(FAR,TAR)at fixed threshold - Vendor K - VISIT\_POE - Right Index

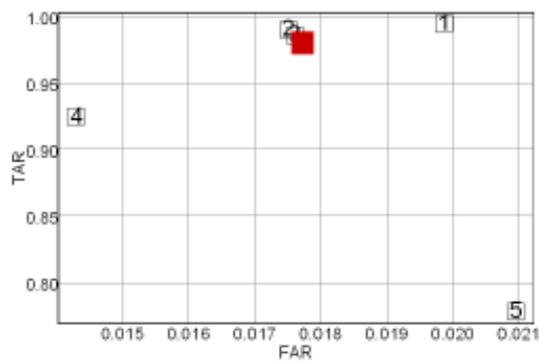
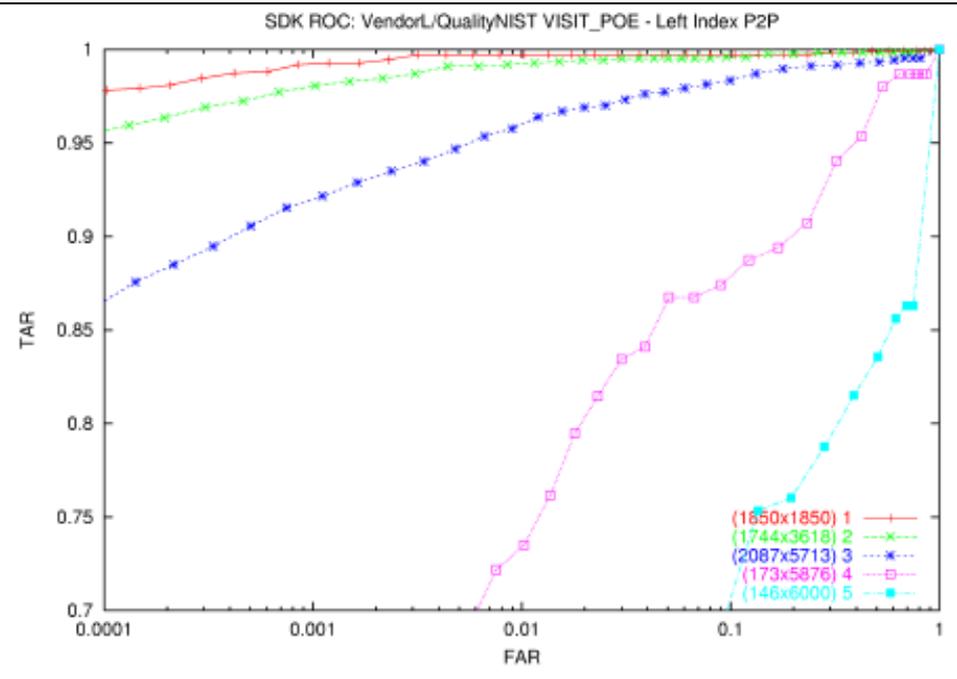


図21-b. ROCに対する品質の影響－ベンダーK－データセットVISIT\_POE－右手人差し指



Vendor L - VISIT POE - Left index threshold=2100 (far_tar)=(0.01,0.952)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0102	0.0151	0.0090	0.0075	0.0065
TAR	0.997	0.994	0.958	0.722	0.514

(FAR,TAR)at fixed threshold - Vendor L - VISIT\_POE - LeftIndex

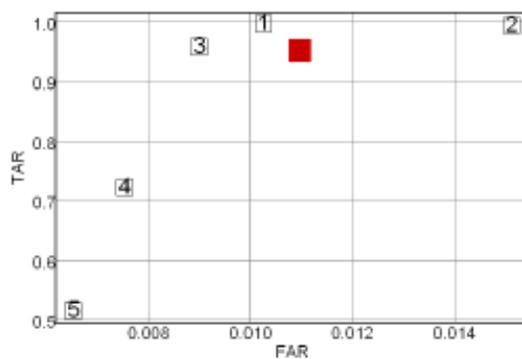
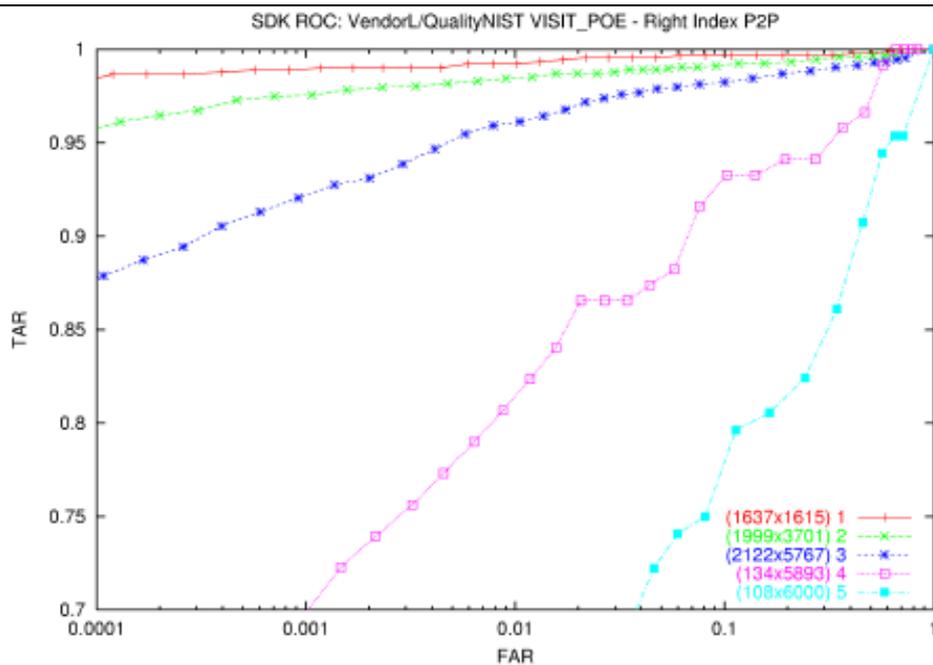


図22-a. ROCに対する品質の影響ーベンダーLーデータセットVISIT\_POEー左手人差し指



Vendor L - VISIT POE - Right index threshold=2000 (far,tar)=(0.012,0.962)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0101	0.0157	0.0104	0.0087	0.0073
TAR	0.992	0.987	0.961	0.807	0.602

(FAR,TAR)at fixed threshold - Vendor L - VISIT\_POE - Right Index

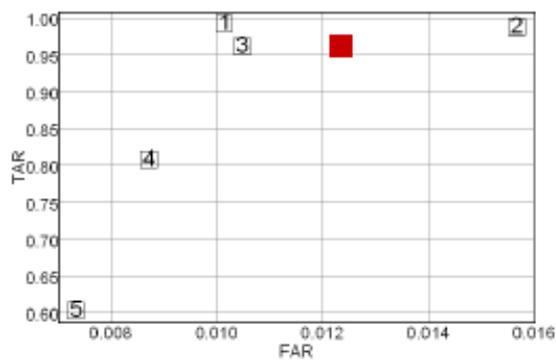
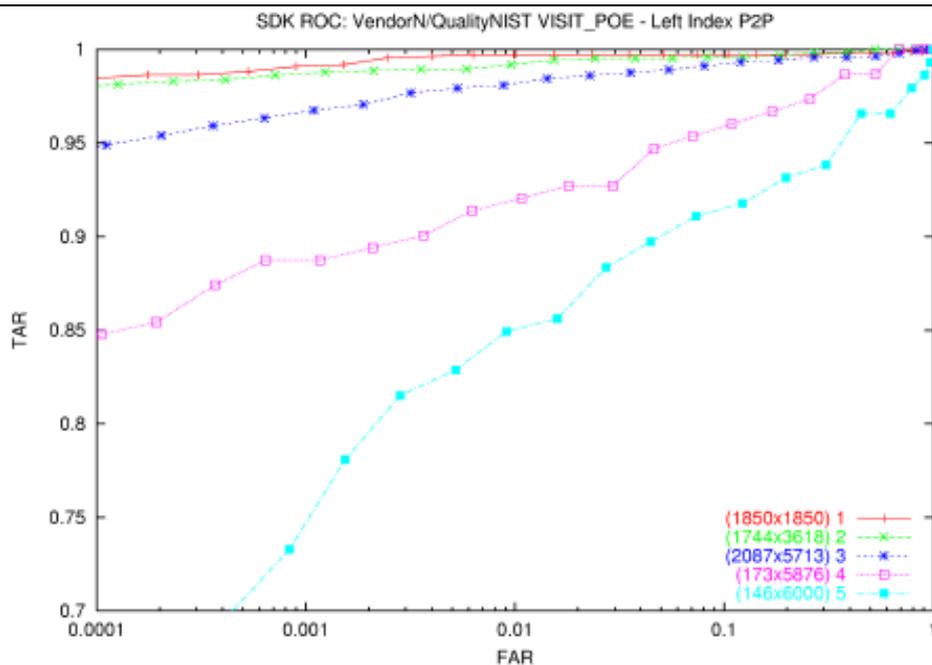


図22-b. ROCに対する品質の影響－ベンダーL－データセットVISIT\_POE－右手人差し指



Vendor N - VISIT POE - Left index					
threshold=36 (far,tar)=(0.01,0.979)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0234	0.0096	0.0088	0.0062	0.0028
TAR	0.997	0.992	0.981	0.914	0.815

(FAR,TAR)at fixed threshold - Vendor N - VISIT\_POE - LeftIndex

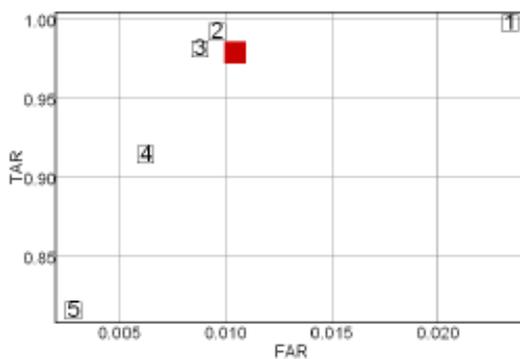
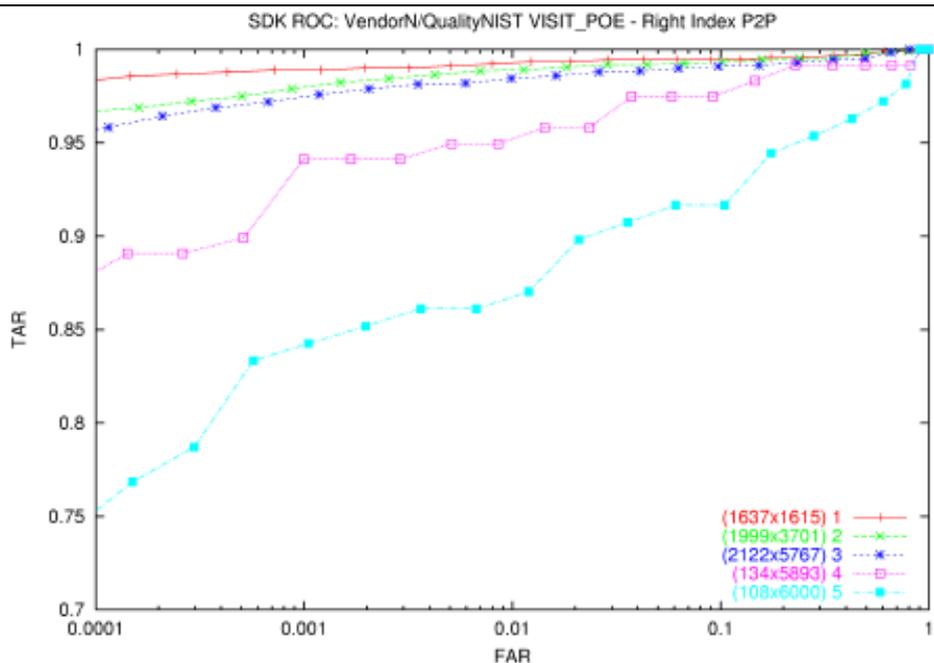


図23-a. ROCに対する品質の影響ーベンダーNーデータセットVISIT\_POEー左手人差し指



Vendor N - VISIT_POE - Right index threshold=34 (far,tar)=(0.01,0.984)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0287	0.0113	0.0099	0.0086	0.0036
TAR	0.994	0.989	0.984	0.949	0.861

(FAR,TAR)at fixed threshold - Vendor N - VISIT\_POE - Right Index

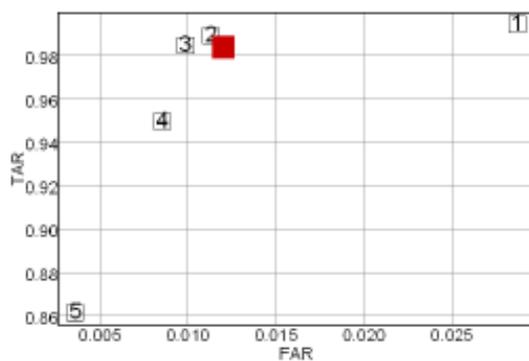
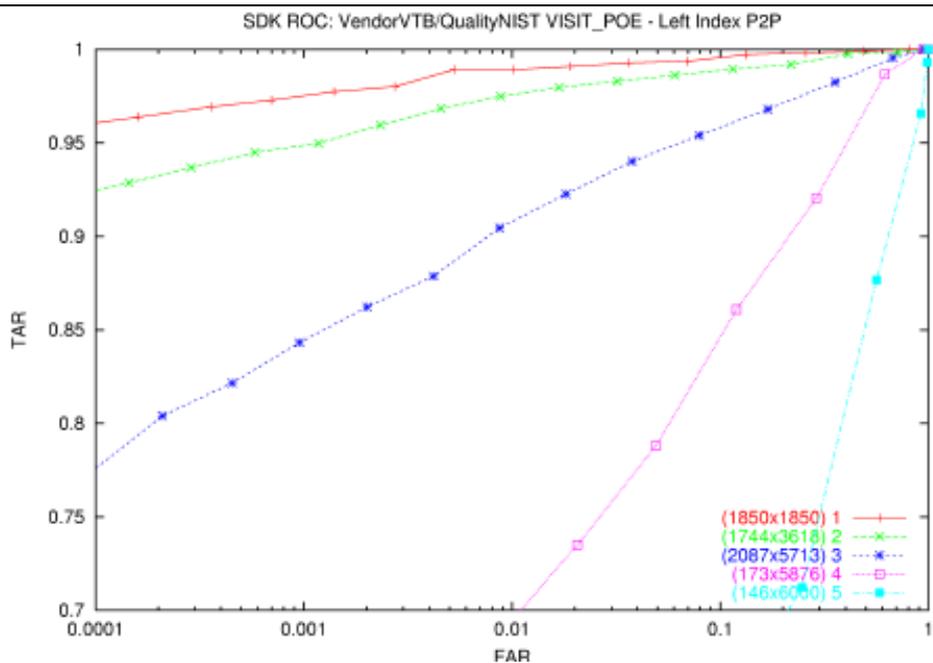


図23-b. ROCに対する品質の影響－ベンダーN－データセットVISIT\_POE－右手人差し指



Vendor VTB – VISIT POE – Left index threshold=18 (far,tar)=(0.011,0.916)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.0191	0.0168	0.0087	0.0040	0.0028
TAR	0.991	0.980	0.905	0.623	0.397

(FAR,TAR)at fixed threshold - Vendor VTB - VISIT\_POE - LeftIndex

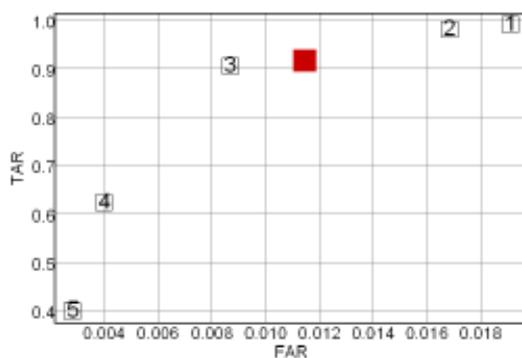
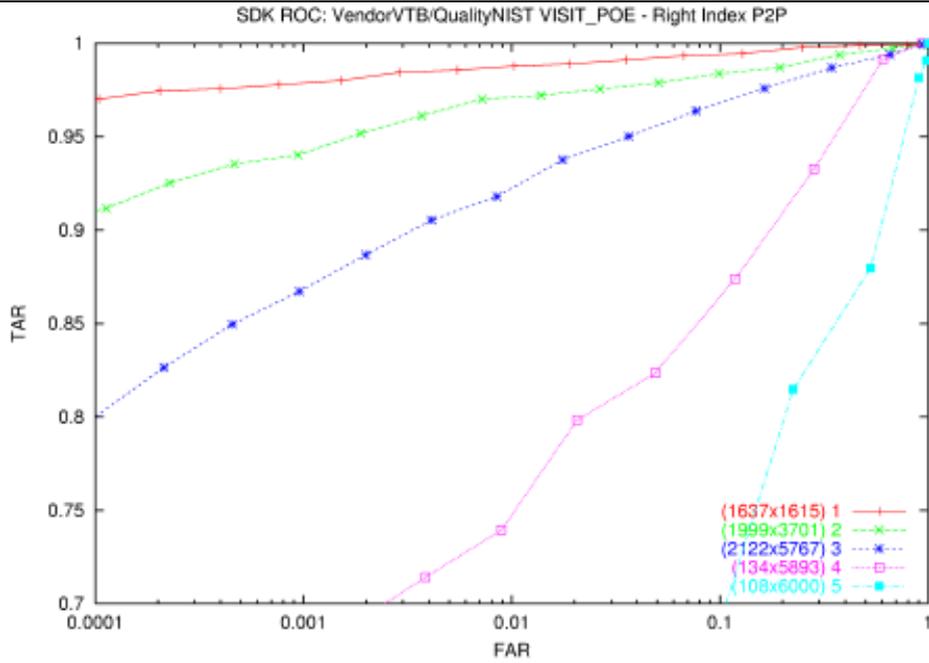


図24-a. ROCに対する品質の影響ーベンダーVTBーデータセットVISIT\_POEー左手人差し指



Vendor VTB – VISIT_POE – Right index threshold=18 (far,tar)=(0.01,0.93)					
Quality	1(excellent)	2(very good)	3(good)	4(fair)	5(poor)
FAR	0.019	0.0138	0.0085	0.0039	0.0020
TAR	0.989	0.972	0.918	0.714	0.435

(FAR,TAR)at fixed threshold - Vendor VTB - VISIT\_POE - Right Index

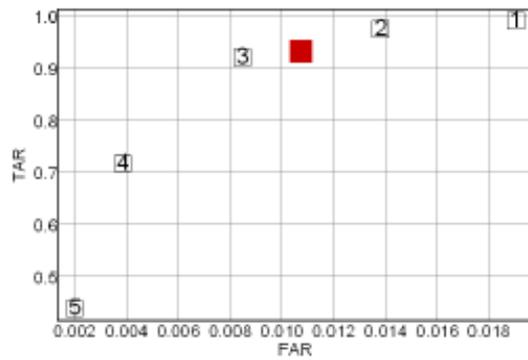


図24-b. ROCに対する品質の影響－ベンダーVTB－データセットVISIT\_POE－右手人差し指

---

## 6 結論

このレポートで、われわれは照合プログラムの性能を予測できる指紋品質の評価方法を策定した。これには、指紋品質の客観的な評価方法が含まれる。ここでの品質の定義は、ほかの種類の実体認証に適用することができ、適切な特徴抽出を行えば、あらゆる形式のバイOMETリックサンプルの品質を評価するために使用できる。

この文書では、NIST NFIS指紋解析ソフトウェアから品質判定用の特徴ベクトルを生成するために使われる方法について説明している。次に、画像品質値を生成する人工ニューラルネットワークのトレーニングに使われる方法を示している。これらの画像品質値を、さまざまな指紋画像データと指紋照合システムの300個の組み合わせについてテストし、すべてのシステムおよびデータセットについて照合性能を予測できることがわかった。US-VISIT POEデータに関するレポートの本文に示されたテスト結果は、この方法がニューラルネットワークのトレーニングに使われなかった照合プログラムとデータの組み合わせに対してもきわめて正確であることを示している。

識別や検証のために特定のバイOMETリックサンプルの品質を自動的に一貫して決定できるかどうかは、広範囲にわたるアプリケーションで問題になっている。品質の低いバイOMETリックサンプルを判定できれば、その情報を使って新しいデータの採取を改善できる。この品質測定基準を使って、品質の悪いバイOMETリックサンプルを品質のよいサンプルに置き換えることで、長期保管されている実体認証ギャラリーを選択的に改善することもできる。複数の形式が混在する実体認証データに対しては、品質のよいバイOMETリックサンプルが混在データの中心となるように重みを選択することができる。これらのアプリケーションは、いずれも識別や検証の前にバイOMETリックサンプルの品質を決定することを要件としている。ここに提示した画像品質測定基準は、この要件を満たしている。これらのアプリケーションは、そのほとんどがバイOMETリックサンプルの品質をデータの採取中にリアルタイムで計算することを要件としている。ここに示した方法では、この要件を満たすのに十分な速度が得られる。

---

**参考文献**

1. A.Jain, L. Hong, S. Pantanki, R. Bolle. Identity authentication using fingerprints, In Proceedings of the IEEE, Vol. 85, No. 9, pp. 1365-1388. 1997
2. M. Y. Yao, S. Pantanki, N. Haas, N. Ratha, R. Bolle. Quantifying Quality: A case study in fingerprints
3. C. Watson. NIST Special Database 29 Plain and Rolled Images from Paired Fingerprint Cards, NIST, November 2001
4. Friedman, J. H. and Stuetzle, W. (1981). Projection pursuit regression. Journal of the American Statistical Association **76**, 817-823.
5. R. A. Hicklin, H. Korves, B. Ulery, M. Zoepfl, M. Bone, P. Grother, R. Micheals, S. Otto, C. Watson. Fingerprint Vendor Technology Evaluation (FpVTE) 2003. NISTIR 7123
6. C. Watson, A. Wilson, K. Marshall, M. Indovina, R. Snelick. Studies of one-to-one fingerprint matching with vendor SDK matchers. NISTIR 7119
7. M.D. Garris, C.I. Watson, R.M. McCabe, C.L. Wilson. User's guide to NIST fingerprint image software. NISTIR 6813
8. R.M. McCabe. Data format for the interchange of fingerprint, facial, scar mark & tattoo (SMT) information. American national standard ANSI/NIST-ITL 1-2000, July 2000
9. Rev. 12-84. U.S. Department of Justice, Federal Bureau of Investigation. The science of fingerprints.
10. C.L. Wilson, C.I. Watson, M.D. Garris, A. Hicklin. Studies of fingerprint matching using the NIST verification test bed. NISTIR 7020
11. N. Nill and B.H. Bouzas. Objective image quality measure derived from digital image power spectra. Optical Engineering, April 1992. Vol. 31 no. 4 pp.813-825
12. Bolle et al. System and methods for determining the quality of fingerprint images. United States patent number US596356, 1999.
13. L.L. Shen, A. Kot and W.M Koo. Quality measures of fingerprint images. 3rd international conference AVBPA 2001, p182-271, June 2001.
14. N.K. Ratha and R.M. Bolle. Fingerprint image quality estimation. IBM computer science research report RC21622, 1999.
15. L. Hong, Y. Wan and A.K. Jain. Fingerprint image enhancement: algorithm and performance evaluation. IEEE transaction on pattern analysis and machine intelligence, vol. 20, no.8, August 1998.
16. E. Lim, X. Jiang, W. Yau. Fingerprint quality and validity analysis. IEEE ICIP 2002
17. C.L. Wilson, J.L. Blue, O.M. Omidvar. Training Dynamics and Neural Network Performance. Neural Networks, vol. 10, No. 5, pp. 907-923, 1997.
18. J.L. Blue and P.J. Grother. Training feed forward networks using conjugate gradients. NISTIR 4776 and in Conference on Character Recognition and Digitizer technologies, vol. 1661, pp. 179-190, San Jose, February 1992.

- 
19. D. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization, *Mathematical programming B*, vol. 45, 503-528, 1989.
  20. O.M. Omidvar and C. L. Wilson. Information content in neural net optimization, NISTIR 4766, and in *Journal of connection science*, 6:91-103, 1993.
  21. S. Wood, C.L. Wilson. Studies of plain to rolled fingerprint matching using NISY Algorithm Test Bed (ATB). NISTIR 7112.
  22. C. L. Wilson, M.D. Garris, C.I. Watson. Matching performance for the US-VISIT IDENT system using flat fingerprints. NISTIR 7110
  23. Hastie, T. and Tibshirani, R. (1990). *Generalized Additive Models*. Chapman and Hall, London.

## 付録 A

ROC

データセットの品質分布

この付録は、別のファイルに含まれる。

## 付録 B

特徴ベクトルの成分と性能のトレリス図

この付録は、別のファイルに含まれる。